Modified Nucleic Acids: Structural Studies and Applications in Biosensing

By

James Carr-Smith

UNIVERSITY OF BIRMINGHAM

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Chemistry

College of Engineering and Physical Sciences

University of Birmingham

July 2015

Abstract

The modification of natural nucleic acids or synthesis of novel DNA mimics can facilitate new structure, function and properties. In particular, the use of modified nucleic acids for applications in biosensing has become a popular field of study given the desire for rapid and reliable theranostic devices. The aim of the projects detailed in this thesis was to study a range of DNA modifications, with a view to gaining an enhanced understanding of their effects on DNA structure, but also on their ability to act as sensing platforms for the extraction of important biological information stored within DNA targets.

The four projects included in this thesis are as follows;

1. Ferrocene based mimics of deoxyribose nucleoside units were incorporated directly into the backbone of DNA to form conjugates, and their structure tuned to promote stable, hydrogen bonding interactions with natural DNA nucleobases. Furthermore, these FcNAs show remarkable base-pair recognition properties as evidenced by thermal melting studies. Initial electrochemical studies suggest potential for applications in biosensing.
2. The study of FcNA-DNA conjugates bearing thymines and the ability to bind mercury(II) ions. Investigations were carried out using thermal melting studies, circular dichroism and mass spectrometry. The systems show some interesting binding properties that suggest they can be used in the future development of oligonucleotide based-mercury sensors.
3. A rare example of metal centred macrocycle that can be incorporated into the backbone of DNA. A host of spectroscopic and electrochemical techniques were utilised to show the ability of intercalating motifs to transduce information regarding the local genetic environment of the incorporated modification, in an

unprecedented example of a metal macrocycle SNP sensor incorporated into DNA.

4. Finally, the work towards the development of virus bound oligonucleotide primers for biosensing is shown. Full characterisation and binding behaviour of these conjugates was investigated using a range of spectroscopic techniques including fluorescence anisotropy and linear dichroism spectroscopy. The bound primers are shown to take part in an unprecedented example of a PCR reaction taking place on a viral surface. The successful amplification of pathogenic DNA can be monitored by linear dichroism spectroscopy in producing observable changes in particle alignment in flow. This work provides great promise for the future design of rapid, mobile, multi-modal biosensing devices.

I dedicate this thesis to my wonderful parents, wife and daughter.

Acknowledgements

Firstly, I wish to thank my supervisor Prof. Jim Tucker for providing me with the opportunity to carry out my Ph.D research project within his group. His advice and guidance on both scientific and personal matters has been invaluable throughout. I always enjoyed our conversations about Queens Park Rangers (someone has to support them) and the contrasts between them and my beloved Chelsea Football Club. Thanks also for helping with proof-reading and allowing me to be involved both in paper writing and submission processes.

Secondly, thanks must go to the entirety of the Tucker Group. More specifically, thanks to Dr. Jean-Louis Duprey for your support and knowledge on all things DNA, working with you on the macrocycle and ferrocene based work has been really enjoyable. To Huy, Jon and Antoine, thanks for all your hard work in producing the phosphoramidites, I know the demands were always high. Andrea, I am grateful for your help in showing me the ropes of electrochemistry. Thanks to Rosie for always providing a good laugh and a physicist's point-of-view. Haydn, I am so glad you joined the group, you were invaluable in helping with all things bacteriophage related; but more importantly I am so grateful for the excellent friendship I have formed with you and Maddy. Gemma, where do I start? Thanks for everything; for being a truly brilliant best friend, a wonderful godparent to my daughter and my tea drinking buddy. Our early morning beverages always helped through the more difficult days. Without you around this Ph.D would have been so much less enjoyable. Finally, thanks to all the other group members, David, Holly, Jack, Pete and John Z, who have all at some point been around for support and a good laugh.

A huge thank you to Dr. Sarah Horswell, her help and patience in improving my electrochemical understanding has been fantastic. Our chats and discussions have always been enjoyable. Your help with proof-reading is also greatly appreciated.

Thanks to the staff of the analytical facilities at the University of Birmingham, namely Dr. Chi Tsang, Peter Ashton and Dr. Neil Spencer for providing a superb service but also for providing a wealth of information that has helped with so many problems. Thanks also to Dr. Paula Mendes and Dr. Tim Dafforn for allowing me to collaborate with your very helpful groups.

I am grateful to the University of Birmingham, School of Chemistry and the EPRSC for providing me with the funding that has allowed me to complete my research.

Thank you to all my brilliant friends in Birmingham: Mike, Danielle, Haydn, Maddy, Rich, Zadeena, Gemma and my best man Matt. The past few years have been great fun, I hope our Friday nights continue for many years to come. Mike, getting to know you has been incredible, thanks for making me an usher and re-introducing me to hockey. Haydn, my very own personal toolbox, my house would be falling down without you. Matt, thanks for everything you did for the wedding, organising it would have been so much more difficult without your help, and what a brilliant speech and stag do.

Most importantly I must say a huge thanks to my wonderful family. Without your support I certainly wouldn't have had the courage to start this journey 8 years ago. To my Mum and Dad, I cannot put into words how grateful I am for everything you have done for me. You truly are the best parents a son could ask for. Chelsey, Grandma, Grandad, Grandad Horse, Paul, Eartha, Megan and Izzy: thanks for making all our times spent together memorable; the Christmas breaks always a highlight of my year.

Finally, my wonderful wife Jen, thank you for being the rock that has helped me through tough days and made the great ones even more memorable. You brought into the world our wonderful daughter, Poppy, who has such a brilliant little character already. I could not be more grateful for everything you do for both of us. I'm looking forward to many more years of fun and surprises.

Table of Contents

List of Abbreviations

A	Amps/Current
Ab	Abasic
Abs	Absorbance
ampR	Ampicillin resistance gene
AuNP	Gold nanoparticle
b.p.	Base pairs
C	Bulk Concentration
CD	Circular Dichroism
CE	Counter Electrode/Auxiliary Electrode
Cp	Cyclopentadienyl
CPG	Controlled Pore Glass
CV	Cyclic Voltammetry
D	Diffusion coefficient / cm^2 s^{-1}
Da	Daltons
DCM	Dichloromethane
ΔE_p	Peak Separation
ΔH	Enthalpy change
Δi	Difference current
ΔG	Gibbs Free Energy change
ΔS	Entropy change
DIPEA	Diisopropylethylamine
DMSO	Dimethylsulfoxide
DMTr	Dimethoxytrityl
DNA	Deoxyribose nucleic acid
dNTPs	Deoxynucleoside triphosphates
DTT	Dithiolthreitol
E	Potential / V
ELISA	Enzyme Linked Immunosorbent Assay
E_p^a	Anodic peak potential / V
E_p^c	Cathodic peak potential / V

ESI	Electrospray Ionisation
E_λ	Switching potential
F	Faraday Constant
FA	Fluorescence Anistropy
FAM	Fluorescein amidite
FcNA	Ferrocene Nucleic Acid
FRET	Förster Resonance Energy Transfer
Γ	Surface roughness
HOMO	Highest Occupied Molecular Orbital
HPLC	High Performance Liquid Chromatography
ICD	Induced Circular Dichroism
$i_{forward}$	Forward current / A
i_p^a	Anodic peak current / A
i_p^c	Cathodic peak current / A
$i_{reverse}$	Reverse current / A
LD	Linear Dichroism
LUMO	Lowest Unoccupied Molecular Orbital
M	Metal
M13	M13 bacteriophage
MeOH	Methanol
MeCN	Acetonitrile
MS	Mass Spectrometry
n	Number of electrons
NanoPCR	Nanoparticle assisted Polymerase Chain Reaction
NHS	N-Hydroxysuccinimide
nm	Nanometer
NMR	Nuclear Magnetic Resonance
PCR	Polymerase Chain Reaction
PNA	Peptide Nucleic Acid
RET	Resonance Energy Transfer
RE	Reference Electrode
RNA	Ribose nucleic acid
RP-HPLC	Reverse Phase – High Performance Liquid Chromatography
RT	Room Temperature
RT-PCR	Real Time – Polymerase Chain Reaction

Chapter 1

Introduction

1.1 Biosensing

In nature, biomolecule recognition events are ubiquitous due to use in triggering of other processes. These triggers are typically conformational changes, hybridization events and redox changes.[1] Such is the complexity of biological systems, owing to millions of years of evolution that these recognition systems are difficult to reproduce or mimic for applications outside of the physiological conditions afforded by the body. The ability of Nature's "biosensors" to respond and sense minute concentrations of particular analytes in highly complex biological systems has led to vast amounts of research on artificial biosensors for potential applications in medicinal diagnostics and therapeutics, military detection devices and gene expression. Biosensors are defined as sensors based on biological motifs such as an antibody or enzyme, that use a biological mechanism to report information regarding the presence of an analyte or a change in environment conditions.[2]

1.1.1 Theranostics and Point-of-Care Biosensing

The rapid rise in degenerative and hereditary diseases across the globe, particularly so in developed countries, has seen the need for personalised health diagnostic devices increase greatly. The ability to assess an individual's health rapidly and before a disease can develop is a challenge that has led to many scientists developing new diagnostic tools. In particular, genetic screening for characteristic strains of cancers or degenerative diseases such as Alzheimer's, means that there is a growing need for platforms that can distinguish between small differences in genetic material. A number of challenges are associated with point-of-care biosensing, the greatest of which is sensing/extracting molecules of interest from their complex biological mixtures. Although this challenge is not such an issue in some applications, it can be somewhat difficult to overcome. In order for sensors to become "mobile" or point-of-care devices, it is also important that sensors respond rapidly to biomolecular inputs, otherwise the work would have to be carried out in a laboratory where lengthy and often laborious protocols are required. In many applications and industries, being

able to work with mobile sensing devices is much more advantageous. For example, with accurate, user-friendly devices that are run without the need for reagents; a paramedic could make on the scene decisions regarding the diagnosis and treatment of a patient without the need to wait for in-hospital laboratory support. In the drive towards fast point-of-care devices it is important to develop devices/assays that are both user-friendly and provide clear analysis with high reliability. It is also important to consider the development of personalised medicine that could be prescribed *ad hoc* according to a given analyte concentration which would improve the care of a patient based on their individual need.

1.1.2 Immunochemistry and the Pitfalls

In recent years, immunochemistry has been utilized for sensing an array of different biomolecules. However this approach is limited in that antibody binding of antigens does not produce easily measurable signals such as electron release, light emission or even measurable conformational changes.[3] Antibodies are also limited in that they are restricted by *in vivo* parameters; meaning operation of antibodies is not feasible under non-physiological conditions. Furthermore, the lifetimes of antibodies are extremely short and are highly prone to denaturation upon small temperature fluctuations, making storage and hence mobility more challenging.[4] It is due to these factors that, although antibody sensors have been ground-breaking in many examples such as the now widely available pregnancy test,[5] there are still challenges hindering their widespread adoption for point-of-call devices. Therefore it remains the case that many analytical approaches to detection and diagnosis rely upon cumbersome multi-step protocols involving a series of separation and characterization methods, which are not only time consuming but also extremely costly. Naturally available binding biomolecules fall into one of two classes of biosensors.[1] One being interaction with an analyte triggering a binding-induced conformational change of the binding-site that in turn induces an output signal. However, it is the second class of sensor that holds greater potential. This class involves the functionalising of biomolecules so that

they respond to external inputs such as the presence of an analyte. It is these biosensors that have become increasingly popular and include biomolecules such as proteins, polypeptides and nucleic acids, of which the latter will form the basis of this thesis.[6]

1.2 Nucleic Acids

Nucleic acids are arguably the single most important building block for the existence of life. Their structure and range of different functions in a cell cycle are crucial in the development and growth of living organisms. Deoxyribose nucleic acid (DNA) contains the underlying genetic information of all living organisms, the vast majority of which is stored in the nucleus of a cell. Ribose nucleic acid (RNA) has many functions including the transcription of genetic information from DNA in the cell nucleus (mRNA), translation of genetic information into proteins (tRNA) and the catalysis or regulation of biological reactions (ribozymes). Without these coherent and complementary functions, the encoding of important proteins would not be possible and is therefore paramount to all living things.

1.2.1 DNA

Deoxyribose Nucleic Acid (DNA) can be described as a perfectly optimised supramolecular molecule. Over billions of years of evolution, DNA has driven the growth and evolution of single celled organisms to animals, plants and insects. Such is the design and specificity of DNA that it stores all the genetic information required for the growth and reproduction of living matter. The human genome consists of 3 billion base-pairs codes for 100,000 different proteins.[7] It achieves this through the complementary pairing of four aromatic heterocycles; two purines (adenine (A) and guanine (G)) and two pyrimidines (thymine (T) and cytosine (C) (Figure 1.1). These nucleobases are each appended to a deoxy-ribose sugar (Figure 1.2) with each nucleosidic unit separated and conjoined by a phosphodiester unit. It is the combination of these three major functionalities; nucleobase, deoxy-ribose sugar and phosphate linkage that give the molecular structure of DNA.

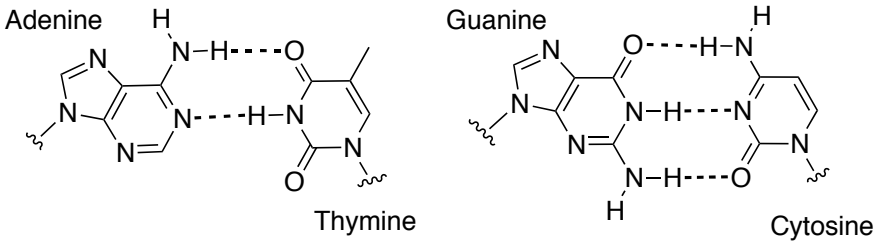


Figure 1.1 - Complementary canonical nucleobases of DNA.

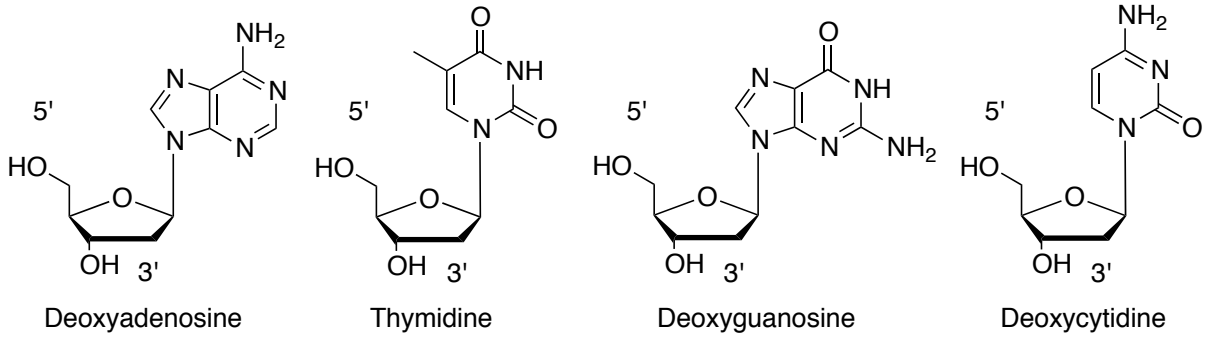


Figure 1.2 - The four canonical nucleosides of DNA; Adenosine; thymidine; guanosine and cytidine.

1.2.2 The Structure of DNA

The structure of DNA was postulated for many years before its elucidation by James D. Watson and Francis Crick in 1953.[8] Their work, for which they were subsequently awarded a Nobel Prize in Physiology or Medicine in 1962 along with Maurice Wilkins, detailed the exact secondary structure of the DNA polymer and it's subunits,[9] a double helical structure which is driven by the complementary base pairing of G-C and A-T base pairs, to form three and two complementary hydrogen bonds respectively. It is the planar aromatic nature of these nucleobase pairs that induces a stacking effect *via* π-π interactions with neighbouring base pairs in the polymer. It was noted that the bond length between C-G and A-T base pairs was equal, which in turn induces a smooth external negatively charged phosphate backbone (shown in Figure 1.3).[8]

The stability of DNA duplexes is dependent on a number of factors, but is influenced mainly by the stacking contributions of the nucleobases. The full extent of this contribution is still an active area of interest to many research groups and it is not thought to be simply due

to the π stack itself.[7] Hydrophobicity, polarisability and overall stacking surface area are important considerations, whilst the overlap of stacked nucleobases is thought to be the largest driving force.[10,11] Although the hydrogen bond network afforded by nucleobase complementarity is an important driving force to the formation of the stack it's contribution to the overall duplex stability is thought to be not as large.[7] The stability of DNA duplexes with respect to single stranded components can be measured by thermal melting experiments (discussed in Chapter 2.2.1).

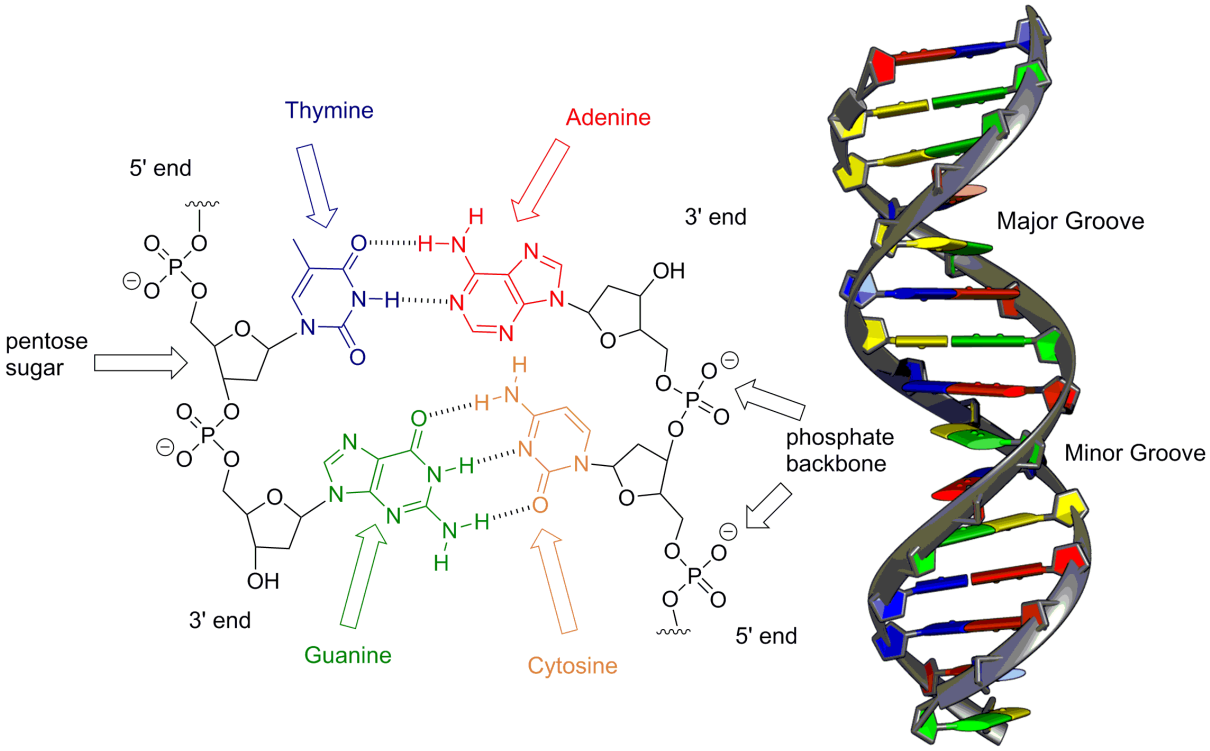


Figure 1.3 - The structure of DNA as elucidated by Watson and Crick. Shown are the complementary base pairs G-C and A-T and the helical structure of B-DNA. Image produced using Chimera.[12]

1.2.3 DNA Conformation

DNA exists in three main conformations, with each having their own unique structural properties. Which one forms depends mainly upon humidity and sequence composition. B-DNA and A-DNA are both right handed helices, with A-DNA having a more compact twist, Z-DNA adopts a left handed helix which is generated in duplexes that involve repeating G-C,

C-G base pairing. The A form is generated when humidity is < 75 %, and is proceeded mostly by B-DNA conformations under natural conditions. A-DNA structures, albeit rare within DNA sequences, are found commonly among RNA duplexes and DNA-RNA hybrids.[13] The structural characteristics of each conformer are shown in Figure 1.4. The addition of DNA binding molecules or the incorporation of modifications can result in perturbations of these structures, but generally the dihedral angles vary rather than the bond-lengths themselves.[13]

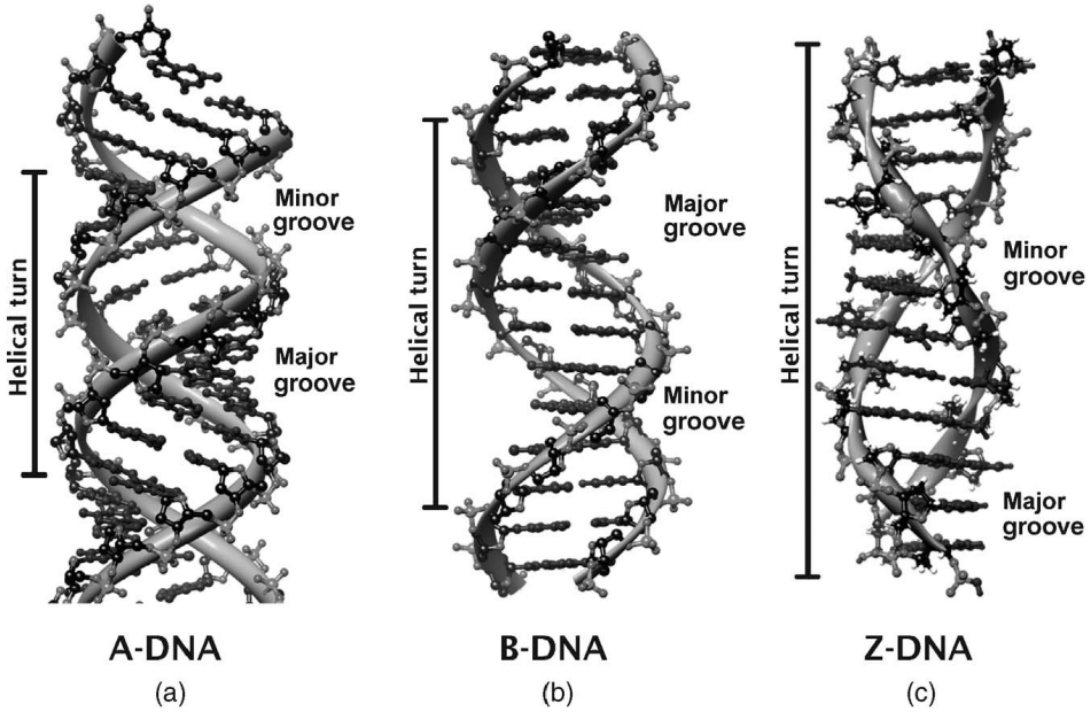


Figure 1.4 Side view of the different DNA conformations a) A-DNA, b) B-DNA and c) Z-DNA. A-DNA has the tightest helical twist and Z-DNA the loosest. Image taken from ref. 13.

1.3 Using DNA for Function

The emergence of automated DNA synthesis in the 1980's meant it was possible to rapidly synthesise nucleic oligonucleotides with exceptional control over sequence and length.[14] Since then, the technology has developed significantly, so much so that synthesizing large quantities of oligonucleotides of up to 100 base pairs long is now unchallenging. The ease of control over the nucleobase sequence has led to the use of DNA oligonucleotides in a wide

range of different applications including nano-electronics, biosensing platforms and supramolecular chemistry. Of great interest in recent years has been the emergence of the field of DNA origami in which any shape, artistic or otherwise, can be designed (Figure 1.5a).[15] The use of DNA origami for applications in therapeutics has led to the creation of DNA boxes that can be opened and closed using highly specific oligonucleotide keys (Figure 1.5b)[16] and new structures that can be used to transport drugs across cell membranes.[17]

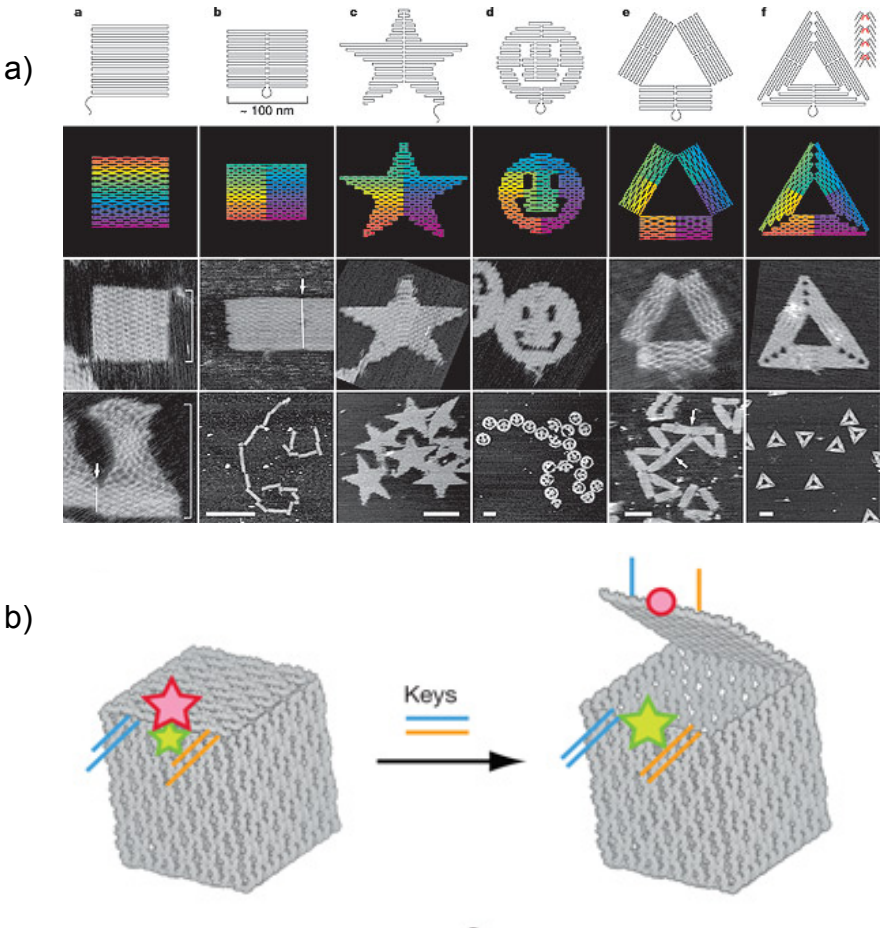


Figure 1.5 - DNA origami nanostructures. a) Stapling of viral DNA using synthetic oligonucleotides.[15] b) A DNA box and key. Image from ref. 16.

1.4 Modified DNA

As part of the ever-increasing drive for enhanced functionality in the field of biotechnology and more specifically DNA technology, the need for control over function and structure has never been greater. If we limit our toolbox to the four canonical bases (A, G, C, T), then regardless of the amount of design or optimisation applied, the scope for enhanced functionality remains limited. However by looking beyond the naturally available nucleobases, it is possible to incorporate non-natural monomeric units with particular functions into the backbone of DNA. Incorporating non-natural motifs into DNA oligonucleotides that can be used to control structure or report on conformational changes has gained prominence in nucleic acid research over the past 20 years. The inclusion of moieties into DNA probes relies heavily on the design of reactive groups that enable the formation of covalent bonds with DNA components. Generally modifications consist of three main functionalities; a reactive group; a spacer group and a signalling or functional group (Figure 1.6).[18]

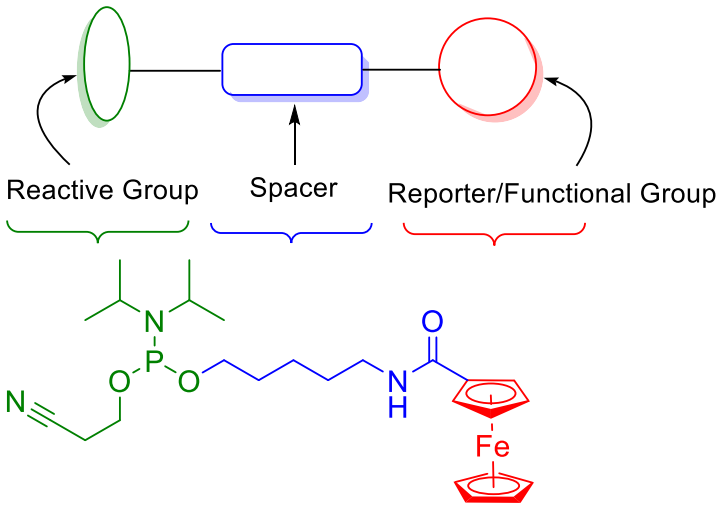
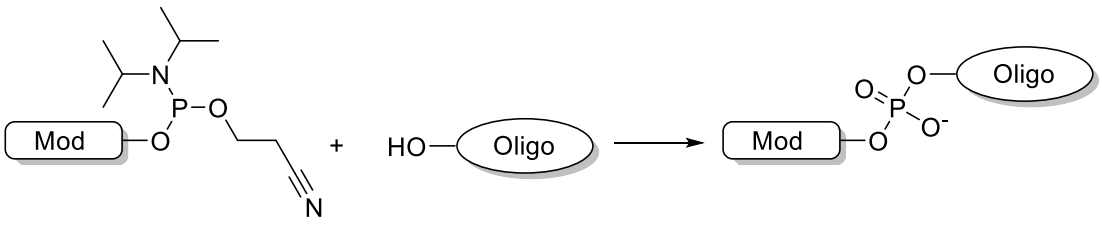


Figure 1.6 - Diagrammatic representation of the main functionalities required for DNA modification. A ferrocene derivative is also shown as an example.

The options for incorporating modifications into oligonucleotides are:

1.4.1 Automated Synthesis

Automated synthesis involves the functionalisation of primary alcohols on sub-units with a reactive group (usually a phosphoramidite) for incorporation *via* oligonucleotide synthesis (discussed in greater detail in Chapter 2.1). In general the overall yields for adding modifications *via* this method are higher than other methods, although greater consideration must be given to the chemical stability of the monomers.

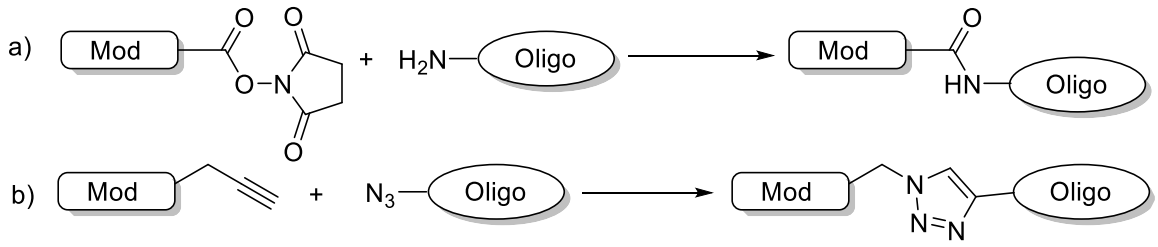


Scheme 1.1 - Incorporation of a modification via phosphoramidite chemistry.

1.4.2 Post Synthetic Labelling

Chemical Ligation

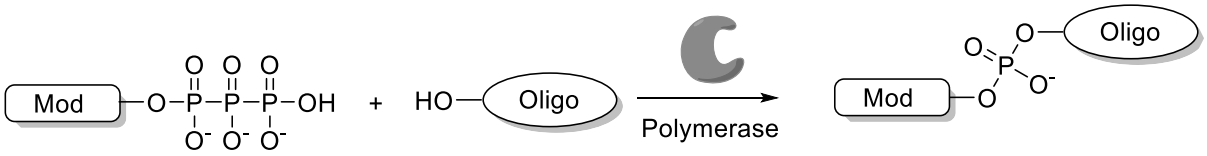
Post oligonucleotide synthesis labelling involves the reaction of a group, on a synthesised oligonucleotide with a compound containing an appropriate modification. Examples include NHS-ester activation (of a primary alcohol) of a modification and subsequent reaction with an amino-labelled oligonucleotide (Scheme 1.2a). Since the term 'Click' chemistry was coined by Sharpless,[19] so called clickable reactions, which are by nature high yielding, low byproduct producing reactions, such as Cu-catalysed alkyne-azide Huisgen cycloadditions[20,21] have become increasingly popular as a tool for conjugating motifs to biomolecules.[18,22] This involves a reactive azide forming a cyclic triazole linkage with an alkyne (Scheme 1.2b). Post synthetic labelling methods generally give smaller overall yields than automated synthesis and usually require additional purification steps and water soluble reagents.



Scheme 1.2 - a) NHS ester catalysed amide bond formation between a modification and oligonucleotide. b) "Click" ligation of a modification to an oligonucleotide.

Enzymatic Incorporation

An alternative method to those mentioned above is the use of a triphosphate moiety *via* enzyme-catalysed ligation (Scheme 1.3). However, synthesising triphosphate monomers is challenging and the method is useful only on a small scale. Nevertheless, it is a popular technique when multiple and sequential incorporations into a strand of DNA are required.



Scheme 1.3 - Enzymatic incorporation of triphosphate moieties into DNA.

1.4.3 Types of Modification

Creating new DNA functionality can be achieved using a whole host of different moieties. Most modifications are used for generating signal outputs or controlling structure. The most common modification types are described below:

1. **Fluorophores** – the use of organic planar aromatic modifications for generating a fluorescence output upon photo-excitation is now the most commonly and widely employed approach, that many tags are now commercially available as ready-made phosphoramidites for automated DNA synthesis. Fluorescence can be used in many areas that include monitoring FRET profiles and anisotropic changes upon binding of DNA targets and analytes.

2. **Redox active motifs** – groups that can readily undergo reversible electron transfer (e.g ferrocene) have been used in DNA modifications since the 1990's for applications in biosensing.[23] There are now a number of metallo-complexes, organometallic and organic motifs that have been added to the DNA backbone and investigated for biosensing and nanoelectronic applications.

3. **Structure manipulation** – Aside from biosensing, the desire to control DNA structure and function through conformational changes has led to the development of many different motifs. For example azobenzene and anthracene functionalities have been utilised to control duplex stability upon exposure to light.[24,25] Also of interest is the use of metal binding ligands for controlling structural and electronic properties.[26,27]

4. **Tethers and Conjugation groups** – Tethering DNA motifs to surfaces for sensing and nanoelectronic applications requires the use of appropriate tether groups such as alkane thiols (gold and platinum). Biotinylated DNA has also been used to form strong non-covalent interactions with streptavidin labelled surfaces. Additionally this type of modification, when incorporated during oligonucleotide synthesis, can enable post synthetic labelling *via* methods such as those discussed in 1.4.2.

1.5 Conclusions

The design of rapid, easy to use biosensing platforms remains a challenge. The synthetic freedom and flexibility of DNA and the emergence of DNA synthesisers allow for ready modification of oligonucleotides for sensing applications. Designing new labels that can mimic and or replace natural nucleotides relies on controlled design of molecular structure and electronic properties of the monomeric units. Furthermore, understanding the cause of changes in signal output afforded by these labels relies on developing a greater understanding of structural and conformational changes and the effects that different oligonucleotide targets impose.

While biosensing remains important from a diagnostics point of view, using supramolecular functionality to control DNA structure holds great promise for a whole host of applications.

1.6 Thesis Outline

This thesis details the work of four different projects carried out during this Ph.D. The nature of each project is that they are distinct from each other in both their context and content. For this reason this thesis is structured with each chapter containing an introduction to the current literature and principles of each project followed by the results and discussion.

Chapter 2 outlines the key techniques and instrumentation used throughout the multiple projects discussed in this thesis. The use of automated oligonucleotide synthesis and electrochemical and spectroscopic methods for the analysis of modified oligonucleotides is detailed.

Chapter 3 discusses the incorporation of a range of ferrocene derivatives (FcNAs) that mimic the structure of naturally occurring nucleic acids. Evidence of their ability to form complementary hydrogen bonding motifs with DNA, RNA and PNA is shown by thermal melting studies alongside circular dichroism spectroscopy. The redox active properties of

these FcNA motifs are probed by electrochemical methodologies to investigate their potential for biosensing applications.

Chapter 4 investigates the ability of FcNA motifs bearing thymines to participate in mercury binding with natural DNA thymines. Optical spectroscopy techniques including thermal melting studies and CD are used to probe the nature of binding events and establish scope for possible biosensing platforms using FcNAs.

Chapter 5 explores the incorporation of metal centred macrocycles into the backbone of synthesised oligonucleotides and their potential use for single nucleotide polymorphism detection. This chapter rationalises the structural effects of these incorporations *via* circular dichroism spectroscopy, thermal melting studies and electrochemical methods.

Chapter 6 details investigations into the use of fibrous virus particles, M13 bacteriophage, as scaffolds for templating DNA amplification *via* the polymerase chain reaction. These systems are studied by shear flow Linear Dichroism spectroscopy for their validity as assays for in-the-field biodetection platforms.

1.7 References

1. A. Vallee-Belisle and K. W. Plaxco, *Curr. Opin. Struct. Biol*, 2010, **20**, 518-526.
2. D. R. Thévenot, K. Toth, R. A. Durst and G. S. Wilson, *Biosens. Bioelectron.*, 2001, **16**, 121-131.
3. A. A. Lubin and K. W. Plaxco, *Acc. Chem. Res.*, 2010, **43**, 496-505.
4. Z. M. Wilner. I, *Aptamers in Bioanalysis*, 1st edn., Wiley, 2008.
5. J. I. Vaitukai, G. T. Ross and G. D. Braunste, *Am. J. Obstet. Gynecol.*, 1972, **113**, 751-758.
6. K. W. Plaxco and H. T. Soh, *Trends. Biotechnol.*, 2011, **29**, 1-5.
7. E. T. Kool, J. C. Morales and K. M. Guckian, *Angew. Chem. Int. Ed.*, 2000, **39**, 990-1009.
8. J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737-738.
9. F. H. C. Crick, J. D. Watson and M. H. F. Wilkins, *The Nobel Prize in Physiology or Medicine 1962*, http://www.nobelprize.org/nobel_prizes/medicine/laureates/1962/, Accessed 20th April, 2015.
10. Y.-P. Pang, J. L. Miller and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 1717-1725.
11. K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, D. C. Tahmassebi and E. T. Kool, *J. Am. Chem. Soc.*, 2000, **122**, 2213-2222.
12. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
13. K. Nakamoto, M. Tsuboi and G. D. Strahan, *Drug-DNA Interactions: Structures and Spectra*, Wiley, 2008.
14. M. D. Matteucci and M. H. Caruthers, *J. Am. Chem. Soc.*, 1981, **103**, 3185-3191.
15. P. W. K. Rothemund, *Nature*, 2006, **440**, 297-302.
16. E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf and J. Kjems, *Nature*, 2009, **459**, 73-76.
17. Q. Jiang, C. Song, J. Nangreave, X. Liu, L. Lin, D. Qiu, Z.-G. Wang, G. Zou, X. Liang, H. Yan and B. Ding, *J. Am. Chem. Soc.*, 2012, **134**, 13396-13403.
18. T. Brown and T. Brown (Jr), *Nucleic Acids Book*, http://www.adtbio.com/nucleic-acids-book, Accessed 09/04/2015, 2015.
19. H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004-2021.
20. C. W. Tornøe, C. Christensen and M. Meldal, *The Journal of Organic Chemistry*, 2002, **67**, 3057-3064.

21. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596-2599.
22. K. Horisawa, *Front. Physiol.*, 2014, **5**, 1-6.
23. R. C. Mucic, M. K. Herrlein, C. A. Mirkin and R. L. Letsinger, *Chem. Commun.*, 1996, 555-557.
24. J. Manchester, D. M. Bassani, J.-L. H. A. Duprey, L. Giordano, J. S. Vyle, Z.-y. Zhao and J. H. R. Tucker, *J. Am. Chem. Soc.*, 2012, **134**, 10791-10794.
25. H. Asanuma, X. Liang, T. Yoshida and M. Komiyama, *ChemBioChem*, 2001, **2**, 39-44.
26. G. H. Clever, C. Kaul and T. Carell, *Angew. Chem. Int. Ed.*, 2007, **46**, 6226-6236.
27. G. H. Clever and M. Shionoya, *Coord. Chem. Rev.*, 2010, **254**, 2391-2402.

SMCC	Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate
SNP	Single Nucleotide Polymorphism
SWV	Square Wave Voltammetry
TAMRA	Carboxytetramethyl rhodamine
TCEP	Tris(2-carboxyethyl)phosphine
THF	Tetrahydrofuran
T_m	Thermal melting temperature
TOF	Time-of-Flight
UV-vis	Ultraviolet-visible
V	Voltage
vs.	Versus
wtM13	Wild type M13 bacteriophage
WE	Working Electrode

DNA/RNA Terms:

A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil

Chapter 2

Techniques

This chapter details the techniques used across the various chapters of this thesis.

2.1 Oligonucleotide Synthesis

In the 1950's chemists began to develop methodologies to efficiently create linkages between individual nucleoside units. Colin B. Reese has extensively reviewed the four main approaches.[1] The most commonly employed method today is the phosphoramidite method pioneered by Marvin Caruthers and co-workers in 1981.[2,3] Their work utilised a reactive 3'-O-(*N,N*-diisopropyl phosphoramidite) substituent appended to the 3' OH of the individual nucleosides. In order to reduce side reactions and increase coupling efficiencies, the other functional groups (nucleobases and 5' OH) are protected using appropriate and cleavable groups (Figure 2.1). The dimethoxytrityl (DMT) group protects the 5' OH from taking part in side reactions during synthetic steps and is readily cleaved under acidic conditions. The most common base protecting groups are shown below in Figure 2.1 and are all removed under basic conditions (usually a 30% ammonia solution).

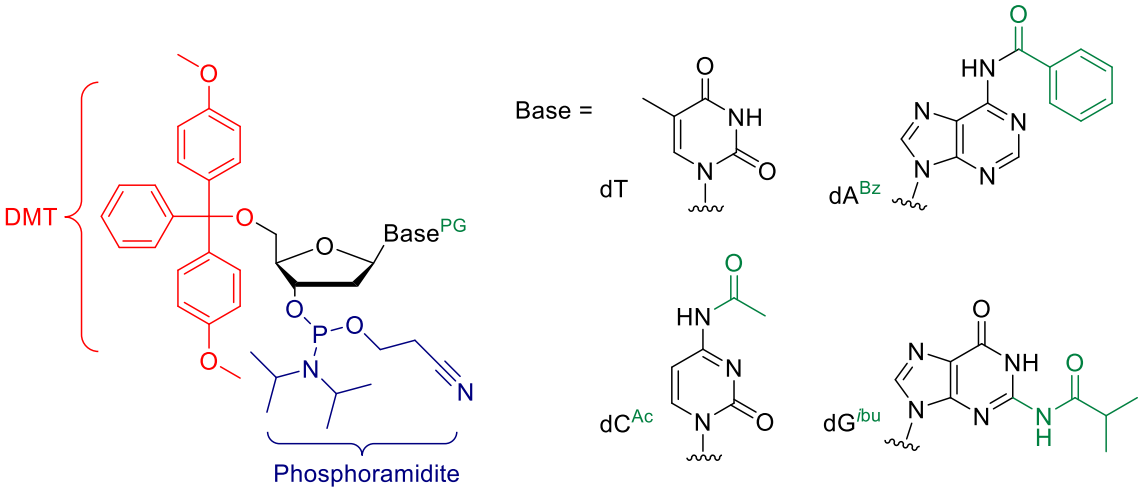


Figure 2.1 - Protected deoxy-nucleoside units for automated oligonucleotide synthesis.

Oligonucleotide synthesis *via* this method involves the stepwise addition of individual mononucleotide phosphoramidite units to a solid-support bound nucleotide unit. The use of a solid-support increases both coupling yields and allows for the maximum possible recovery of synthesised products. The solid-support columns are designed so that they fit readily to a

DNA synthesiser which controls the stepwise addition of the multiple chemical reagents required for synthesis to take place. The emergence of this type of technology was vital in improving coupling efficiencies and also the length of oligonucleotide sequences that could be created.[4]

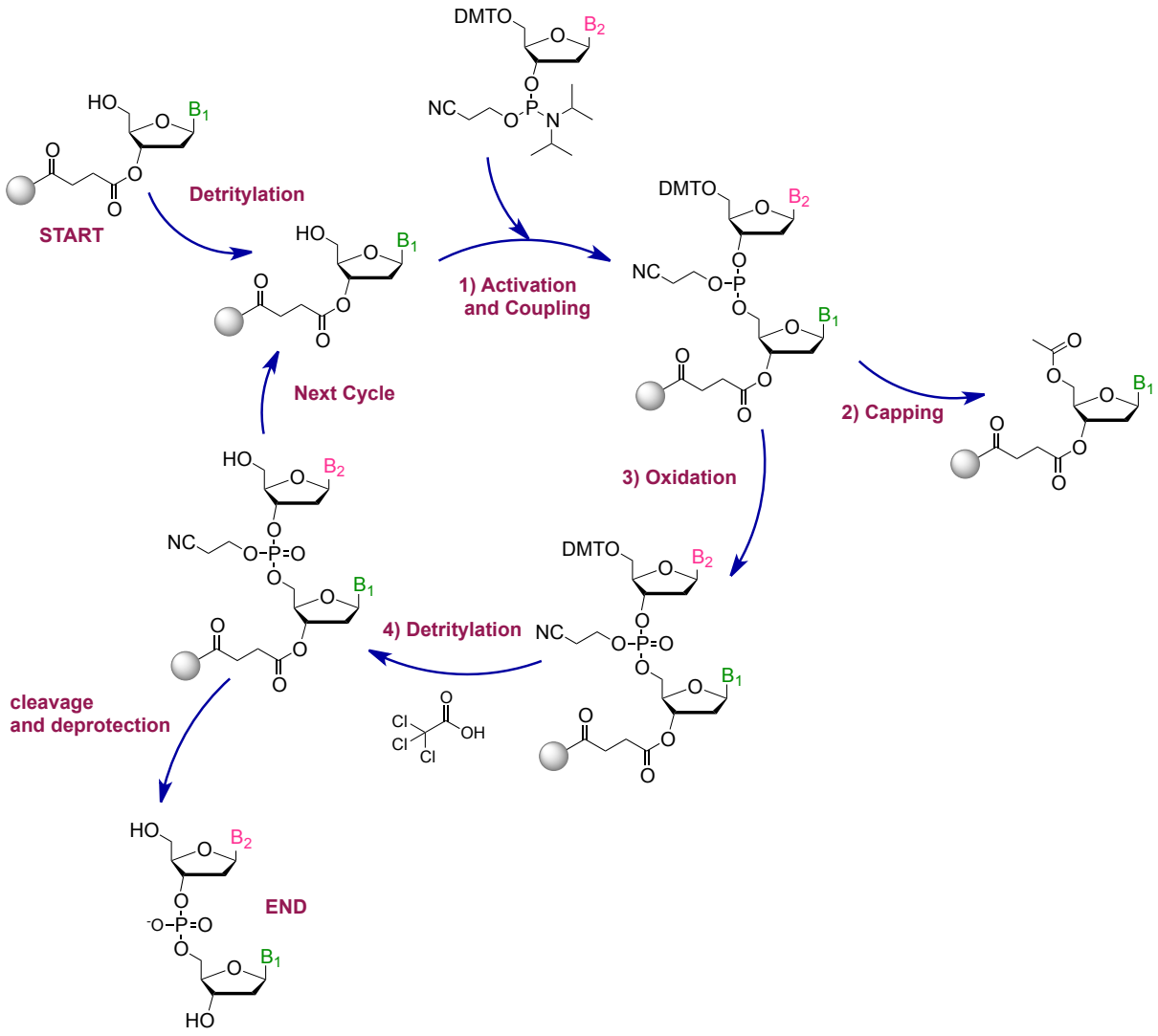


Figure 2.2 - Automated, solid-phase oligonucleotide synthesis cycle utilizing the phosphoramidite method.

The oligonucleotide synthesis cycle can be considered to consist of four main steps (shown in Figure 2.2):[5]

1. **Activation and Coupling -** Prior to the first coupling, the solid support bound nucleotide is treated with trichloroacetic acid in order to remove the 5' OH protecting DMT group to produce a free hydroxyl functionality at the 5' end (Figure 2.3). Positively charged trityl cations are used throughout the synthesis to measure the coupling efficiencies of each step by virtue of its UV-vis absorbance.

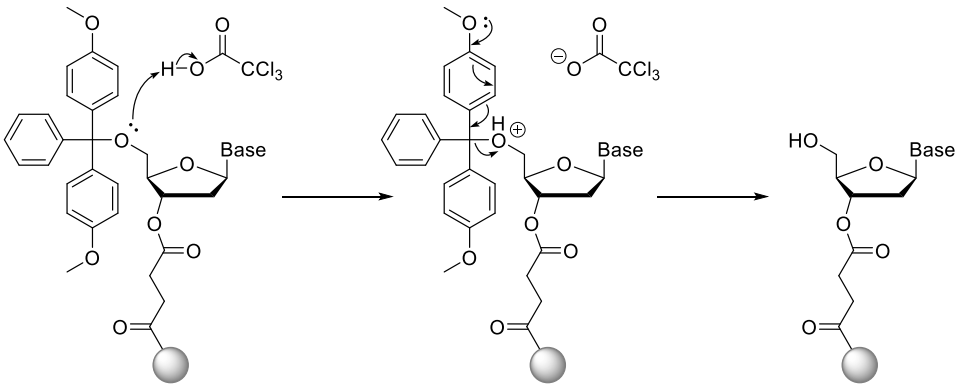


Figure 2.3 - Detritylation of DMT protected 5' OH with trichloroacetic acid.

Following detritylation, the support bound nucleotide reacts with the incoming nucleoside phosphoramidite of choice, which is added in large excess in combination with an activator, generally tetrazole. These reagents are delivered in acetonitrile. The activator generates an excellent leaving group by protonating the phosphoramidite which is subsequently displaced by the reactive support bound 5' hydroxyl, which attacks the phosphorous atom (Figure 2.4). This reaction generates a phosphite triester linkage between the two nucleosides.[6]

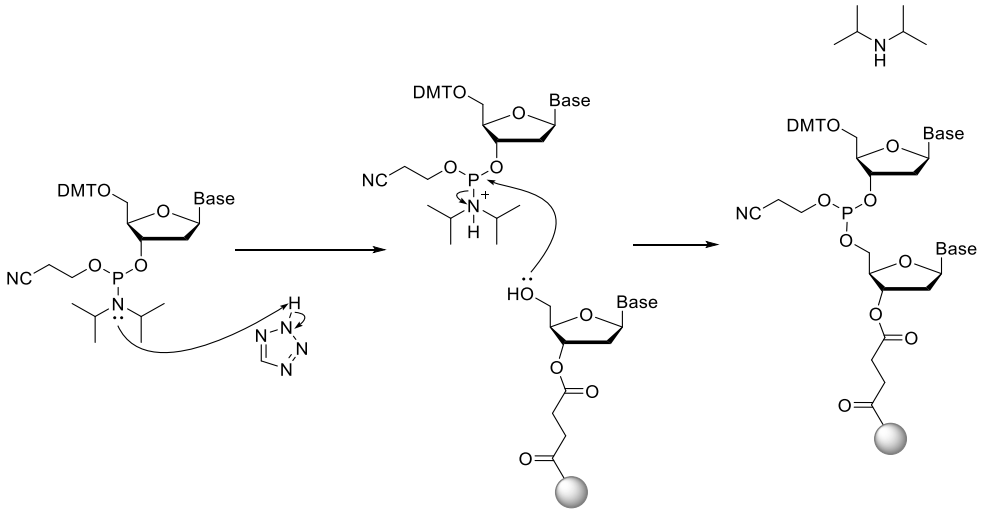


Figure 2.4 - Activation and coupling of nucleoside phosphoramidite.

2. **Capping –** Although yields for individual coupling steps can often reach 99.5%, even the most efficient processes result in a notable loss of overall yield. The uncoupled 5' OH groups bound to the resin are 'capped' to ensure they are unable to participate in future coupling steps (Figure 2.5). This step ensures that the desired final oligonucleotide is more easily separated from unsuccessful truncated sequences. The capping solution consists of a mixture of acetic anhydride and N-methylimidazole dissolved in THF, which contains small amounts of pyridine. This reaction mixture acetylates the free hydroxyl groups, with the pyridine ensuring that the pH remains basic enough so that the DMT protecting group is not cleaved.

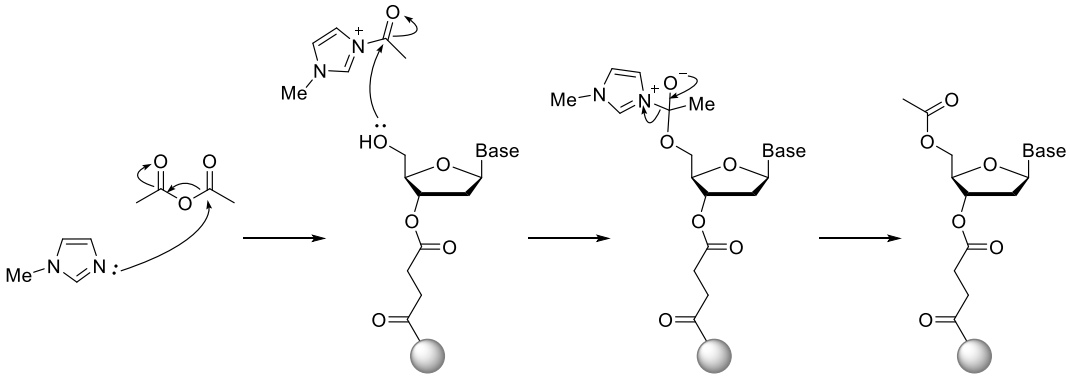


Figure 2.5 – Acetylation (capping) of unreacted 5' OH groups.

3. **Oxidation -** The oxidation step is used to convert the unstable phosphite triester formed in step 1 from P(III) to a more stable P(V) species (Figure 2.6). A solution of iodine in water and pyridine is used to achieve this. The resulting phosphotriester is essentially what is found in the DNA backbone with a cyanoethyl protecting group.

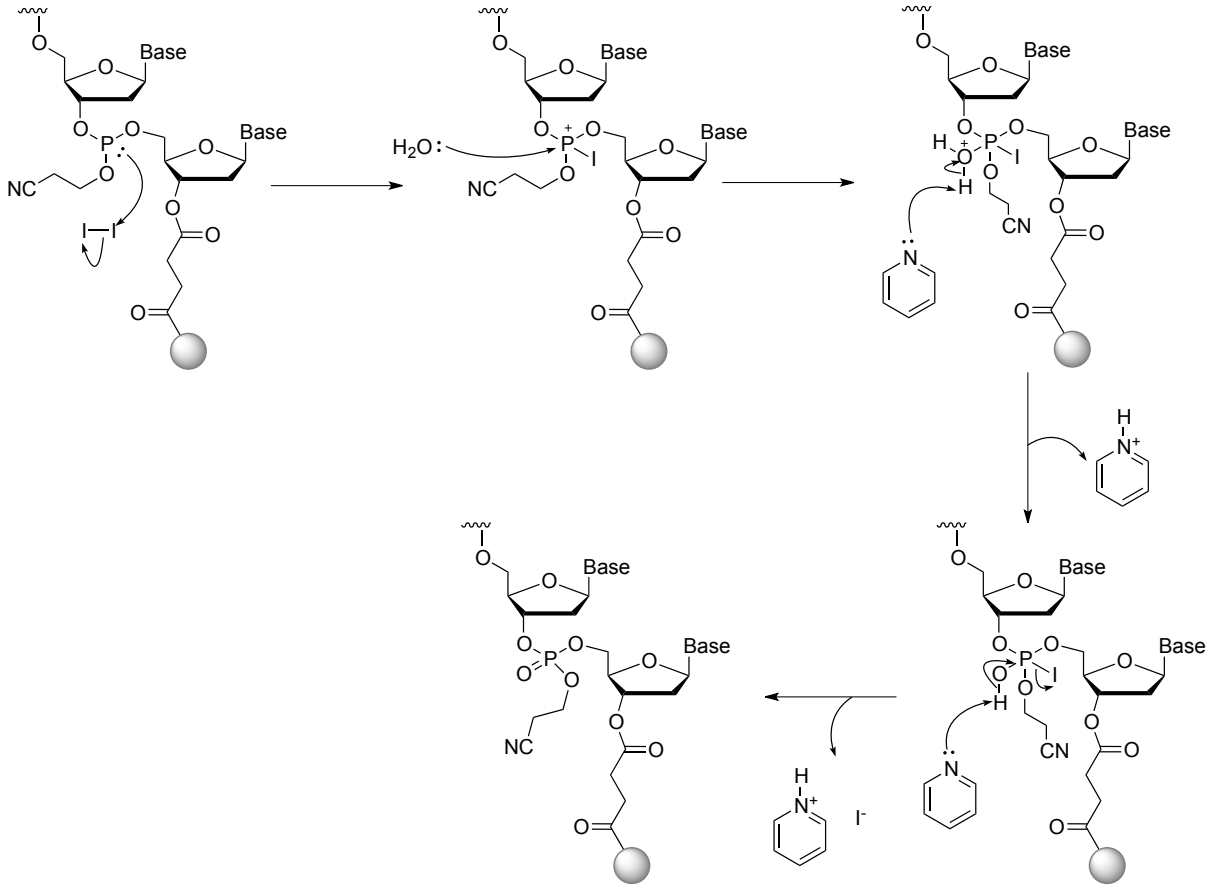


Figure 2.6 - Oxidation of the phosphite triester using iodine and pyridine.

4. **Detritylation -** The coupled nucleosides are now ready to be extended. In order for this to occur, the DMT protecting group on the now coupled nucleoside must be removed in order for the next incoming phosphoramidite to be able to react with it. This step is identical to that prior to the first activation and coupling (see Figure 2.3).

These four steps are repeated an appropriate number of times, introducing the desired nucleoside phosphoramidite during step 1 of each cycle. Once the oligonucleotide sequence

is complete, it is cleaved from the solid support to allow for deprotection of the protected nucleobases (Figure 2.7) and subsequent purification. Generally cleavage is achieved by the addition of 30% ammonia solution. Other less harsh conditions (ULTRA-mild) can be selected for less stable phosphoramidite moieties. Further exposure of the oligonucleotide to 55 °C heat in 30 % ammonia removes the nucleobase protecting groups.

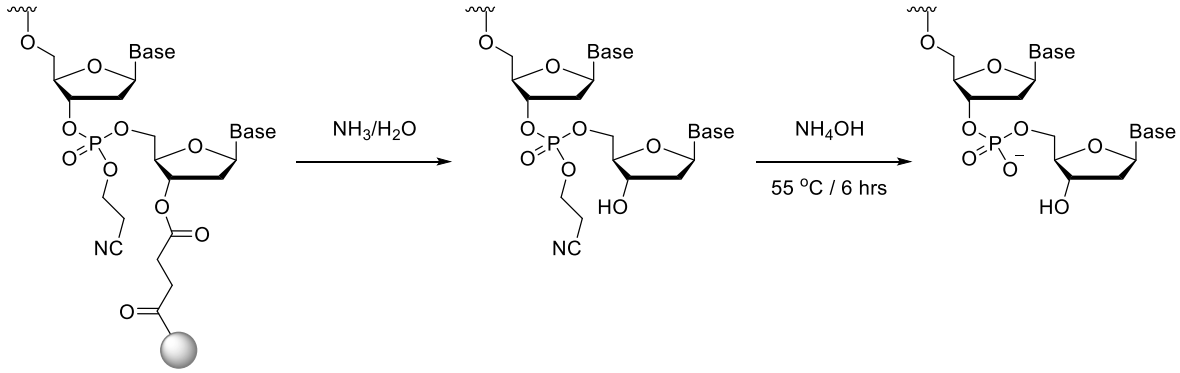


Figure 2.7 - Cleavage and deprotection of the oligonucleotide using aqueous ammonia solutions.

2.1.1 Oligonucleotide Modifications – the phosphoramidite method

Since the invention of this technique and the creation of synthesisers, the phosphoramidite method has been utilised greatly to incorporate a range of chemical constituents into oligonucleotides including:nucleosidic analogs, fluorescent probes, electrochemical tags and conjugation linkages to name a few. In essence, if a monomeric unit bears a primary alcohol it can be phosphitylated and incorporated into oligonucleotides provided it is soluble in the appropriate synthesis solvents and can withstand the harsh conditions of the synthesis (Figure 2.8).

The work in this thesis describes the use of three different modification approaches for uses in biosensing. Each proceeding *via* phosphitylation and subsequent incorporation *via* the method described above. Any changes from the standard coupling methods used are detailed within the text where appropriate.

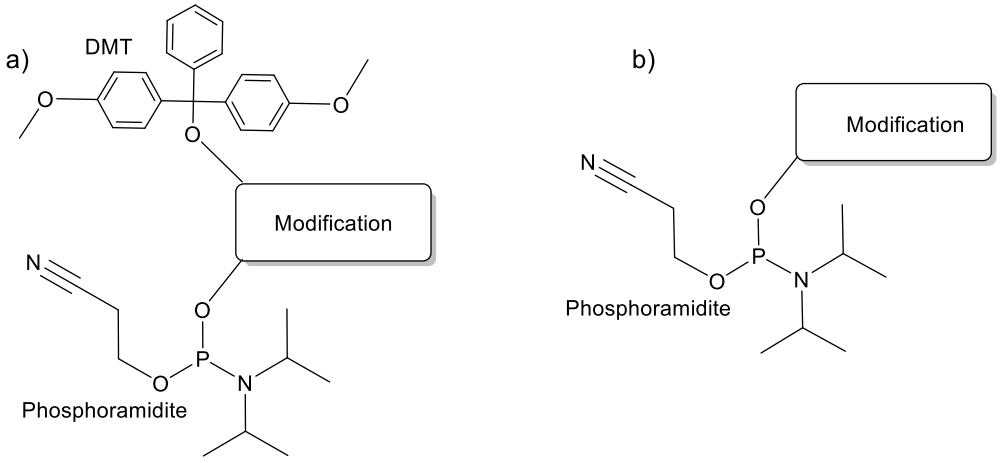


Figure 2.8 - Monomer incorporation via oligonucleotide synthesis. a) Internal and 3' incorporation using DMT and phosphoramidite functionalised diols. b) 5' labelling using a phosphoramidite moiety.

2.2 Optical Spectroscopy for DNA

DNA and its nucleobase heterocycles naturally absorb UV light which allows for the analysis of structural changes through a host of different spectroscopic methods. The absorbance spectrum of DNA in this region is dominated by π to π* transitions of the nucleobases. The two techniques most commonly employed to analyse changes in DNA conformation are detailed below.

2.2.1 Thermal Melting Studies

Monitoring of the UV absorbance spectrum of DNA (usually at 260 nm) as a function of temperature allows for the analysis of the stability of a DNA structure. René Thomas noticed in 1951, that the heating of DNA resulted in large increases in the absorptivity of DNA.[7] The hydrogen bond network that joins complementary base pairs in DNA structures are susceptible to breaking at elevated temperatures and as a result a concomitant loss of the base pairing network and base stack of DNA is lost.[7] The phenomenon of this conformational change and increase in UV absorbance is known as the hyperchromic effect.[7]

This property, utilised alongside a temperature controlled UV-vis spectrometer, can be used to determine the melting temperature (T_m) of DNA structures. A typical thermal melting experiment consists of a sample of DNA being ramped to elevated temperatures (denaturation) and subsequently cooled (renaturation). The UV-vis absorbance at a specific wavelength is plotted as a function of the concentration. The T_m value is measured from the first derivative ($\Delta A/\Delta T^{-1}$) of the thermal melt data (see Figure 2.9).[8] This value can be used to compare the differences in stability between different DNA conformations or sequences. Generally, the T_m of a DNA duplex is dependent on three main factors:

- Length of sequence: the larger the number of complementary base pairs, the greater the stacking contributions and stabilising hydrogen bonds.
- G-C base pairs: the greater the G-C content within a sequence the higher the T_m given the additional hydrogen bond afforded by this pair.
- Salt concentration: the stabilising ability of positive counterions is important in negating the repulsion of the negative phosphate backbone of DNA strands. Typically Na^+ or Mg^{2+} are used.

An extensive review of the analysis of thermal melting curves has been carried out by Mergny and Lacroix.[8]

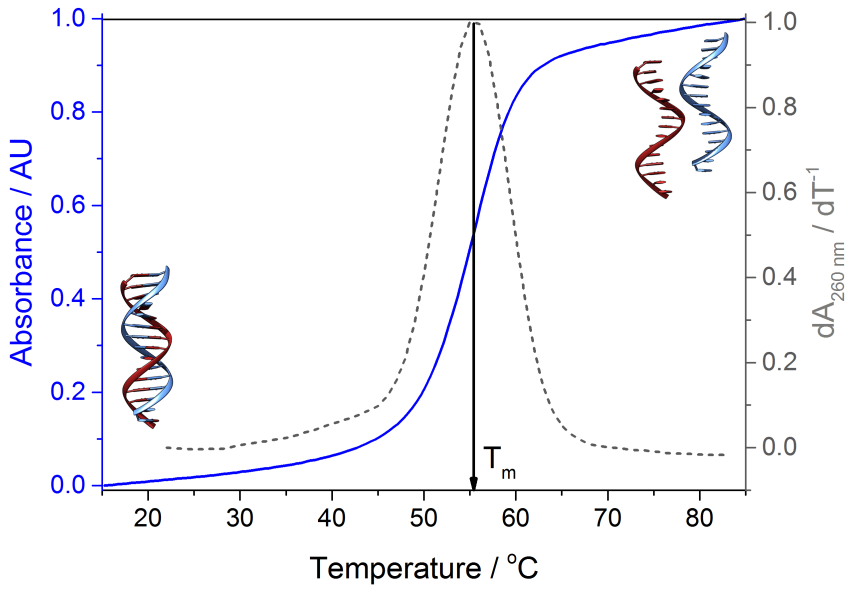


Figure 2.9 - A typical thermal melting curve for duplex DNA showing the measured absorbance increases with temperature (blue solid line) and the first derivative (dashed grey line) from which the T_m is calculated.

This type of experiment can also be applied to the study of protein-DNA interactions; drug-DNA interactions;[9,10] intercalator-DNA interactions;[9] RNA and non-natural nucleic acid structures[11,12] and aptamers.[8]

The T_m values discussed throughout this thesis represent the average of a minimum of three repeats where the T_m quoted is the average of those measurements.

2.2.2 Using T_m Data for Thermodynamic Calculations

Using the T_m values for comparing similar systems can be too restrictive. It fails to inform on the strength of a duplex at a given temperature (i.e. 25°C). A T_m higher than 25°C simply implies that at that temperature the majority of the structure remains hybridised, this value alone cannot be converted into binding constants or used to determine enthalpy and entropy parameters for a system. In fact it could be the case that for two duplexes, both with the same T_m, display very different binding affinities at lower temperatures i.e. those that are generally worked at in a laboratory. In order to obtain a greater understanding of some of the

duplex systems within this thesis, the 'hyperchromicity' method was used to extract thermodynamic parameters for enthalpy (ΔH), entropy (ΔS) and hence Gibb's free energy (ΔG). The method utilises the thermal melting curve data obtained in T_m experiments alongside Van't Hoff plots according to the following assumptions:

The Gibbs free energy is related to the binding affinity by Equation 2.1

Equation 2.1

$$\Delta G = -RT\ ln\ (K_a)$$

Where K_a is the binding affinity and R is the ideal gas constant (8.314 J K^{-1} mol^{-1}). The assumption is made that the system exists in a two-state (bimolecular) equilibrium where molecules are either folded or un-folded.

The first step involves the conversion of the absorbance (A) vs. temperature (T) data (Figure 2.10a) into a fraction of folded/hybridised duplex (θ) (Figure 2.10b). This is achieved by subtraction of the upper and lower baselines of the melting curve. The conversion of absorbance to θ is via Equation 2.2 and results in a profile similar to that in Figure 2.10b.

Equation 2.2

$$\theta_T = (LowerBaseline_T - A_T)/(LowerBaseline_T - UpperBaseline_T)$$

Secondly, the data taken between 0.15 < θ < 0.85, the region where the data is most reliable is converted to an affinity constant, K_a according to Equation 2.3, which describes the conversion for a bimolecular process involving complementary oligonucleotides.

Equation 2.3

$$K_a = \frac{\theta_T}{2C_0(1-\theta_T)^2}$$

Where C_0 is the strand concentration. The data is then plotted as a natural log vs the reciprocal of the temperature (1/T) in kelvin which results in a van't Hoff plot similar to that shown in Figure 2.10c.

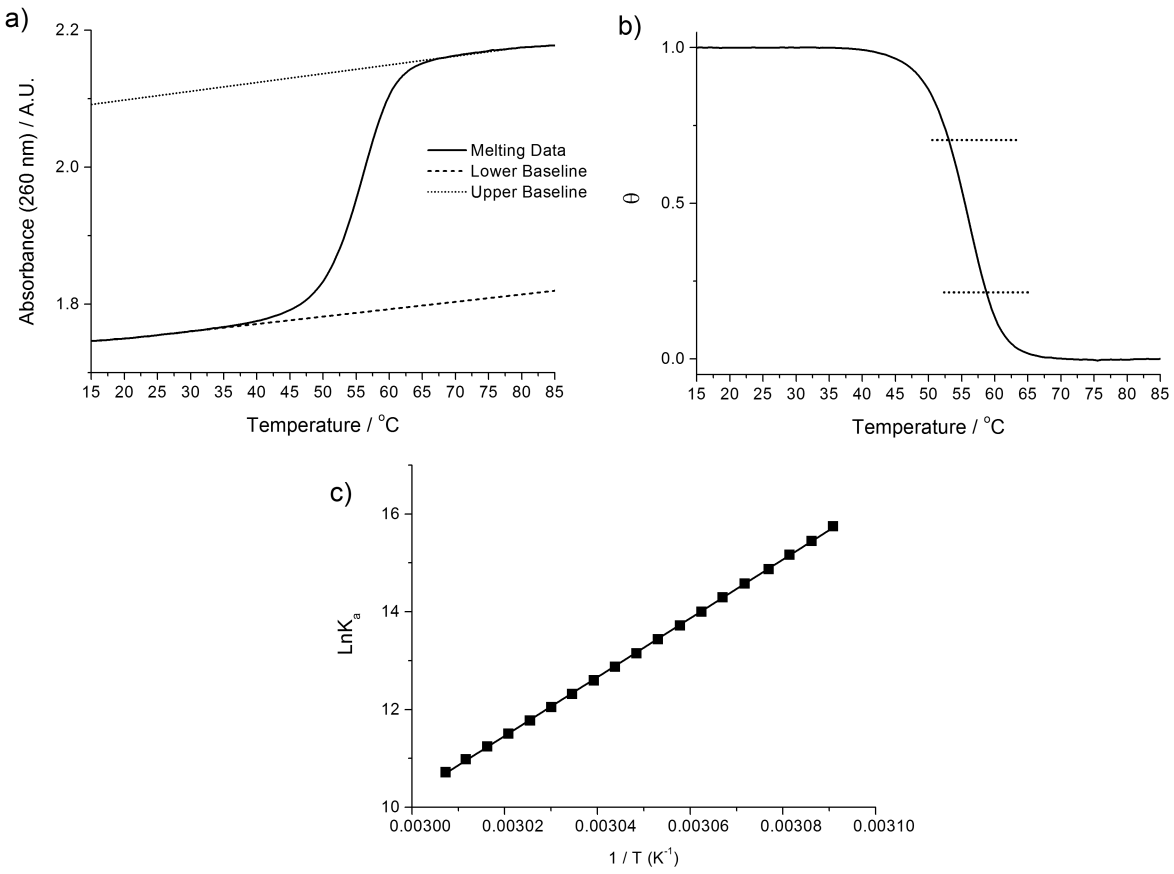


Figure 2.10 - The three stages of determining thermodynamic parameters from thermal melting data. a) Determination of the upper and lower baselines of the melting curve. b) Conversion of melting data to fraction folded (θ). c) Van't Hoff plot of Ln(K_a) of each data point versus 1/T in K. The slope and intercept of the linear plot can be used to determine ΔH and ΔS respectively.

By the relationship between the binding affinity and enthalpy and entropy (shown in Equation 2.4) and the assumption that both ΔH and ΔS are temperature independent the Van't Hoff plot of ln(K_a) vs. 1/T gives a straight line. The slope of that line is defined as – ΔH/R and the Y-intercept is ΔS which is represented by Equation 2.5.

Equation 2.4

$$\Delta G = -RT\ln(K_a) = \Delta H - T\Delta S$$

Equation 2.5

$$\ln(K_a) = \left(-\frac{\Delta H}{R}\right)\frac{1}{T} + \left(\frac{\Delta S}{R}\right)$$

The calculated ΔH value is defined as the Van't Hoff enthalpy and is derived from Equation 2.6 below.

Equation 2.6

$$\Delta H = -R\,\Delta\ln(K_a)\,/\Delta(T^{-1})$$

This method does make some assumptions that do not strictly always hold true but for purposes of drawing comparisons between different systems this is deemed appropriate.[8,13]

2.2.3 Circular Dichroism Spectroscopy

Electromagnetic radiation consists of both an electric and magnetic field that oscillate perpendicular to each other. When an electric field vector maintains constant magnitude in time but plots a helix along the direction of its propagation (k), we observe circularly polarised light. The directionality of the electric field vector and hence helix follows either right or left handedness (Figure 2.11a).[14]

Circular dichroism (CD) is a technique that utilises circularly polarised light to measure the difference in absorbance (ΔA) between right (A_R) and left handed (A_L) light, at a specific wavelength, by an absorbing optical medium (Equation 2.7).

Equation 2.7

$$\Delta A = A_L - A_R$$

Since the helical nature of the polarised light is chiral, the way each form of light interacts with a specific molecule will be different depending on the orientation of the transitions within the molecule. A CD spectropolarimeter shines both right and left handed light on to a sample.

If there is an observed difference between the absorbance of these light directionalities at a given wavelength then an ellipticity (θ) is observed and bands in the CD spectrum produced (Figure 2.11b).

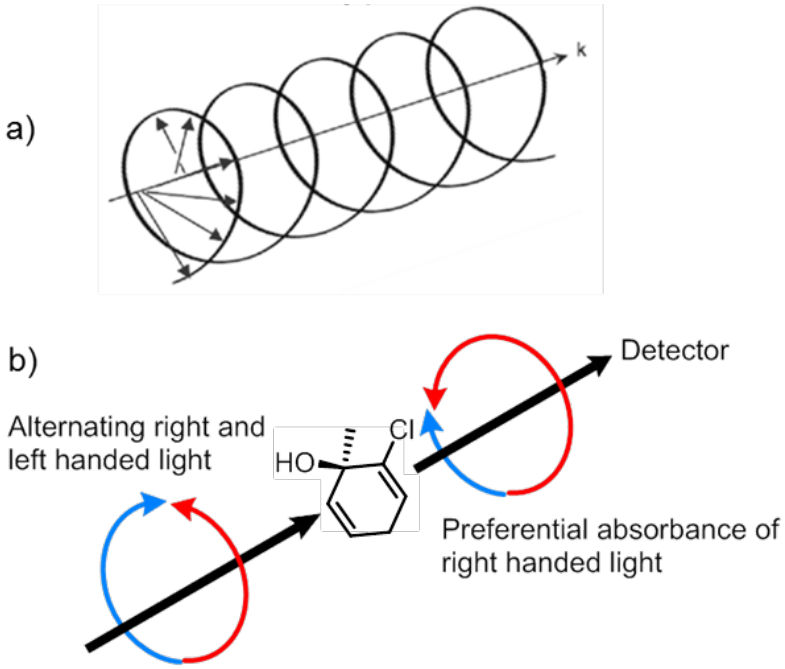


Figure 2.11 - a) Circularly polarised light helix following its propagation direction, k. b) Preferential absorbance of right handed light by chiral molecules.

2.2.4 Circular Dichroism for Nucleic Acids

CD spectroscopy has found many applications across a wide range of research fields including asymmetric catalysis, biochemistry and materials chemistry. Most commonly however, is the use of CD for the analysis of biomolecules such as nucleic acids and proteins.[10,14] The naturally occurring helicity/chirality exhibited by these biomolecules makes CD a particularly attractive tool for understanding their structure. Nucleic acids in particular, as a result of the combination of the base-pair stack, sugar pucker and phosphate backbone display a helical twist (see Chapter 1). The π-π* transitions that dominate the UV-vis spectrum of DNA are alone achiral, but acquire asymmetry by coupling their electronic

transitions with the chiral sugar phosphate backbone. All DNA conformers exhibit a CD maximum typically at 270 nm.

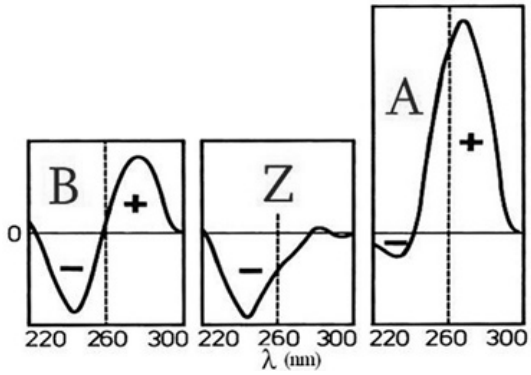


Figure 2.12 - A schematic representation of the typical CD bands observed B, Z and A DNA.[15]

CD is arguably the easiest technique to use in order to acquire information regarding DNA conformation, although analysis of the results is often open to interpretation. DNA polymorphs (A, B and Z) display their own unique CD profiles owing to the differences in base stacking conformations. In particular, and of most interest in this thesis, is the structure of B-DNA which typically displays a CD spectrum consisting of a positive band at 275 nm, negative band at 240 nm and crosses zero at 258 nm. At higher energy wavelengths (*ca.* 220 nm) the sign and magnitude of the bands is dependent mainly on the specific nucleic acid sequence.

It should be noted that the CD results discussed throughout the work in this thesis are used in conjunction with other spectroscopic and electrochemical (*vide infra*) techniques to justify the rationale behind DNA structures.

2.3 Electrochemistry

Electrochemistry is the field of chemistry related to the relationship between electrical and chemical processes. One aspect of this field concerns the chemical changes that take place as a result of a passage of electrical current.[16] This broad field encompasses an understanding of a range of different phenomena (electrophoresis and corrosion) and technologies such as electroplating and batteries.[16] The aspect of electrochemistry relevant to this this thesis is the use of electrochemical methods for understanding redox-active DNA systems.

An electro-active molecule or material can be investigated using techniques that measure electrical outputs as a function of an applied potential, *E*. These electro-active species, upon application of a specific potential, are known to undergo electrochemical changes that result in an exchange of electrons between the species of interest and the electrode, which in turn generates a peak in current. The peaks generated can be used to evaluate redox processes, bulk properties and also for applications in sensing where minute changes can be readily monitored by highly sensitive techniques such as cyclic voltammetry (CV) and square wave voltammetry (SWV). In general, for measuring electro-active species such as DNA, a traditional 3-electrode cell set-up is utilised (shown in Figure 2.13): a working electrode (WE), a counter electrode (CE) and reference electrode (RE), all of which are suspended in an electrolyte solution capable of passing electrical current. These electrodes are connected to a potentiostat, an instrument capable of controlling applied potentials and measuring the electrical outputs as a function of both potential and time. The applied potential is controlled with respect to a reference electrode (often Ag/AgCl or sat. calomel) and the counter electrode is used to collect any current flow generated by the working electrode.

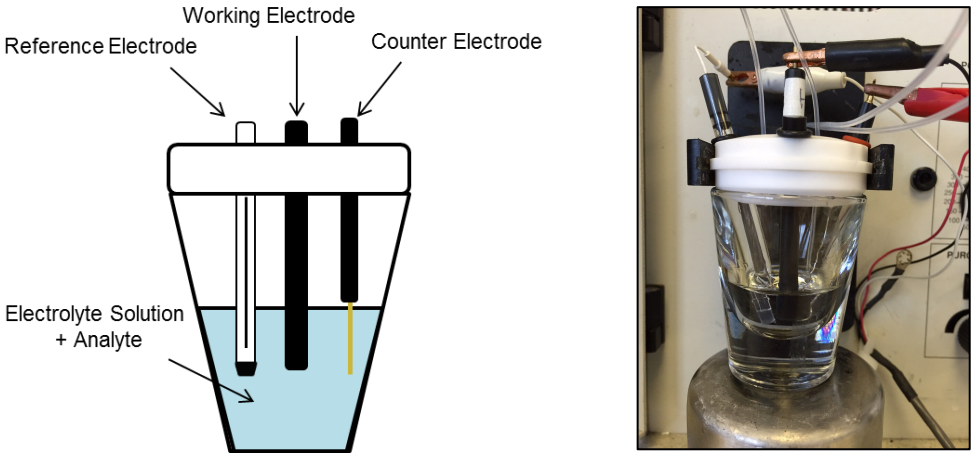


Figure 2.13 - Left - A three electrode cell set-up. Right – BASi cell stand and cell.

Electrochemical methods available for use in electroanalytical applications are in vast supply, however, the two techniques described below form the foundation for the experiments and discussion within this thesis.

2.3.1 Cyclic Voltammetry

Cyclic voltammetry (CV) experiments utilise a linearly sweeped potential that becomes either more positive or more negative as a function of time. After the sweep reaches its switching potential, E_λ, the potential sweep direction is reversed; a schematic is shown in Figure 2.14. During the experiment the potentiostat measures the current flowing in the system and plots this *versus* potential (Figure 2.15). For a system displaying electrochemical reversibility, two peaks are observed, one in the forward scan and one in the reverse for each redox process.[17]

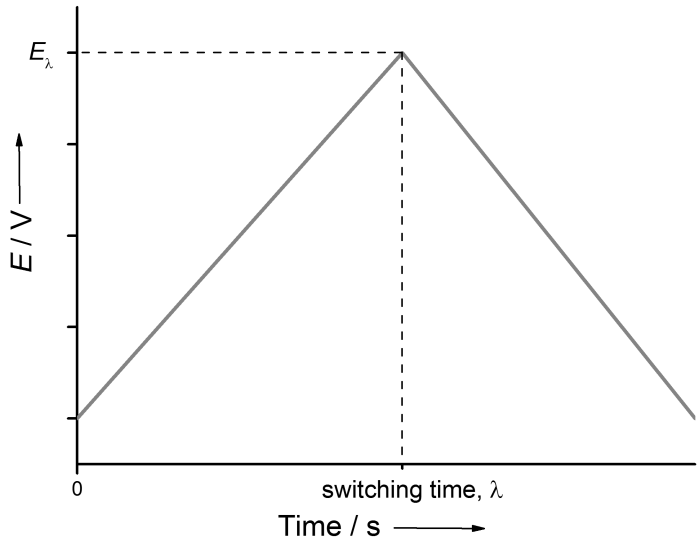


Figure 2.14 - Cyclic potential sweep showing E vs. time.[16,17]

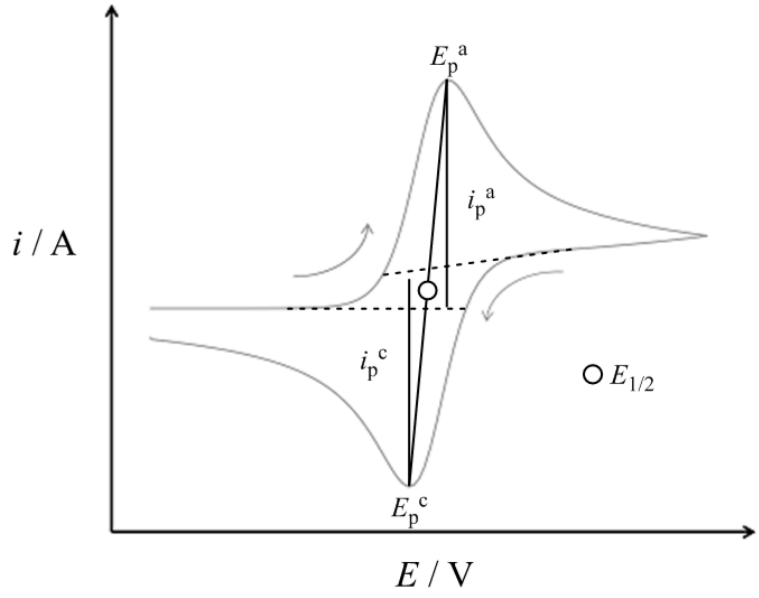


Figure 2.15 - A typical cyclic voltammogram of a reversible redox process. E_p^a = anodic peak potential, E_p^c= cathodic peak potential. i_p^a = anodic peak current, i_p^c = cathodic peak current.[17]

The observed peaks do not display symmetry in relation to each other and do not occur at the same potential. The shape of the voltammogram can be predicted by taking into account kinetic and mass transport effects.[18] From the obtained voltammograms, information

regarding electrochemical parameters can be collected. The $E_{1/2}$ potential, is defined as the $(E_p^a+E_p^c)/2$; where E_p^a = anodic peak potential and E_p^c = cathodic peak potential. Information regarding the electrochemical reversibility is obtained using CV. If i_p^a/i_p^c, the ratio between anodic and cathodic peak currents is equal to 1, and the peak separation ($\Delta E_p = E_p^a - E_p^c$) close to 59 mV for a one electron (n = 1) transfer process, so called Nernstian behaviour, then the system is considered to be reversible. By investigating the current response (i_p^a) to change in scan rate, the time taken to scan from one potential to another, it is possible to determine information on the diffusion rates of the species of interest using the Randles-Sevcik equation (Equation 2.8).[16,19,20] For a reversible electron transfer process in solution, the peak position is invariant with scan rate and the relationship of current is proportional to the square root of the scan rate.

Equation 2.8

$$i_p = 0.4463\, nFAC \left(\frac{nFvD}{RT}\right)^{1/2}$$

Which at 25°C is equal to:

Equation 2.9

$$i_p = 2.69\, x10^5 n^{1/2} A D^{1/2} C v^{1/2}$$

Where:

i_p = peak current (A)

n = number of electrons transferred per molecule

A = surface area of electrode (cm^2)

D = diffusion coefficient of solution species ($cm^2\ s^{-1}$)

C = bulk concentration ($mol\ dm^{-3}$)

v = scan rate (V s^{-1})

A plot of i_p^a vs $\sqrt{v}$ will give a straight line for a quasi-reversible or reversible redox process in solution from which the slope can be used to determine the diffusion coefficient D according to (Equation 2.10).

Equation 2.10

$$Slope = 2.69\times10^5 AD^{1/2}C$$

For a quasi-reversible process, the linear relationship does not hold and peak separation increases with scan rate.

CV was used throughout the work detailed in this thesis to assess, where relevant, the electrochemical behaviour of both single and double stranded oligonucleotide systems.

2.3.2 Square Wave Voltammetry

Pulse voltammetry, and more specifically square wave voltammetry (SWV) provides added benefits over CV for applications in sensing and diagnostics. SWV was invented in 1969 by Krause and Ramaley,[21] it encompasses many of the characteristics of other pulse voltammetry methods; the background/noise suppression of differential pulse voltammetry (DPV) and the diagnostic applications of normal pulse voltammetry (NPV).[21-23] In contrast to CV, SWV utilises a regularly doubly pulsed potential. Its waveform is considered to be that of a staircase, as shown in Figure 2.16.

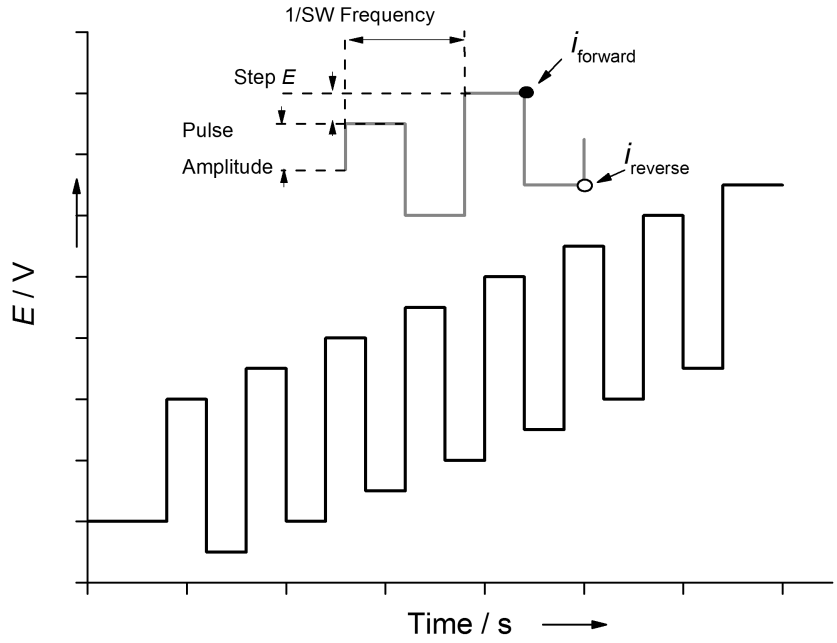


Figure 2.16 - Square wave pulse schematic.[21]

The current is sampled prior to each double pulse ($i_{forward}$) and again before its end ($i_{reverse}$); deduction of the reverse from the forward results in a difference current, Δ*i* (Equation 2.11)*, which* is observed as a symmetrical peak devoid of much of the background signal usually associated with other techniques. An example of an SWV is given in Figure 2.17.

Equation 2.11

$$\Delta i = i_{forward} - i_{reverse}$$

This has the advantage of reducing detection limits and enhancing the sensitivity of electrochemical detection methods.[23] This technique was used throughout the electrochemical sensing studies with oligonucleotide systems discussed in this thesis. Specifically, SWV is utilised for determining the changes in electrochemical signal upon addition of DNA targets.

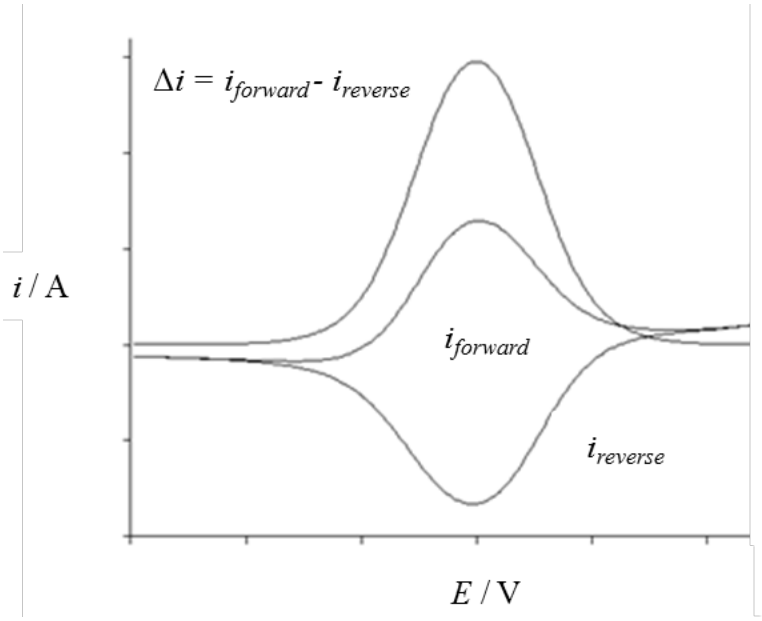


Figure 2.17 - A typical square wave voltammogram.[21,22]

2.4 References

1. C. B. Reese, *Org. Biomol. Chem.*, 2005, **3**, 3851-3868.
2. S. L. Beaucage and M. H. Caruthers, *Tetrahedron Lett.*, 1981, **22**, 1859-1862.
3. M. D. Matteucci and M. H. Caruthers, *J. Am. Chem. Soc.*, 1981, **103**, 3185-3191.
4. M. J. Gait, *Oligonucleotide synthesis : a practical approach*, IRL, Oxford, 1984.
5. T. Brown and T. Brown (Jr), Nucleic Acids Book, www.adtbio.com/nucleic-acids-book, Accessed 09/04/2015, 2015.
6. S. Berner, K. Mühlegger and H. Seliger, *Nucleic Acids Res.*, 1989, **17**, 853-864.
7. R. Thomas, *Gene*, 1993, **135**, 77-79.
8. J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.
9. A. Guédin, L. Lacroix and J.-L. Mergny, in *Drug-DNA Interaction Protocols*, ed. K. R. Fox, Humana Press, 2010, vol. 613, ch. 2, pp. 25-35.
10. K. Nakamoto, M. Tsuboi and G. D. Strahan, *Drug-DNA Interactions: Structures and Spectra*, Wiley, 2008.
11. K. Murayama, Y. Tanaka, T. Toda, H. Kashida and H. Asanuma, *Chemistry*, 2013, **19**, 14151-14158.
12. A. Sen and P. E. Nielsen, *Biophys. J.*, 2006, **90**, 1329-1337.
13. A. Sen and P. E. Nielsen, *Nucleic Acids Res.*, 2007, **35**, 3367-3374.
14. B. Nordén, A. Rodger and T. Dafforn, *Linear Dichroism and Circular Dichroism*, Royal Society of Chemistry, 2010.
15. V. I. Dodero, Z. B. Quirolo and M. A. Sequeira, *Front. Biosci.*, 2011, **16**, 61-73.
16. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, Wiley, 2000.
17. P. T. Kissinger and W. R. Heineman, *J. Chem. Educ.*, 1983, **60**, 702-706.
18. J. A. V. Butler, *J. Chem. Soc., Faraday Trans.*, 1924, **19**, 729-733.
19. J. E. B. Randles, *J. Chem. Soc., Faraday Trans.*, 1948, **44**, 322-327.
20. A. Ševčík, *Collect. Czech. Chem. Commun.*, 1948, **13**, 349-377.
21. L. Ramaley and M. S. Krause, *Anal. Chem.*, 1969, **41**, 1362-1365.
22. J. G. Osteryoung and R. A. Osteryoung, *Anal. Chem.*, 1985, **57**, 101-110.
23. J. G. Osteryoung and J. J. Odea, *J. Electroanal. Chem.*, 1986, **14**, 209-308.

Chapter 3

Ferrocene Nucleic Acid: A DNA Mimic

The work presented in this chapter has been supported by work with other Tucker Group members. Dr. Huy Van Nguyen, Dr. Antoine Sallastrau and Jonathon Kedge synthesised the monomer ferrocene units. Dr. Jean-Louis Duprey aided the structural studies with molecular dynamic simulations and modelling. Additionally, further studies of DNA•PNA systems were supported by Prof. Peter. E. Nielsen and Dr. Anjana Sen at the University of Copenhagen, Denmark.

3.1 Introduction

Naturally occuring nucleic acids DNA and RNA have evolved over time to be the perfect scaffolds for the storage of information, and their structure allows for propagation of that information through specific base-pair interactions during cell replication. How these scaffolds evolved over billions of years is open for debate and intepretation.[1] Most DNA and RNA sequences display similar physiochemical properties, which in contrast to proteins, where a single amino acid change can cause a drastic change in property, could be considered restrictive in terms of their potential use in many applications. However by expanding the toolbox available through modification of nucleic acids, it is possible to introduce new function, interesting new structural and physiochemical properties but also an enhanced understanding of the evolution of life.[1,2] The desire for new materials and novel biocompatible therapeutics has made this endeavour ever more important.

3.1.1 XNA's

Over the past several decades, the synthesis of many DNA mimics has been described. A summary of many of these structures is shown in Figure 3.2. This mimicry usually involves the replacement of one or more of the three repeating units (sugar, phosphate, base) of DNA/RNA; these mimics are often described as *xeno*-nucleic acids (XNA's). Locked nucleic acid (**LNA**), where the ribose sugar unit of RNA is conformationally locked by a bridging CH_2 linkage between a 2' OH and the 5' carbon, has been shown to be particularly stable against enzyme degradation pathways, resulting in its use across a wide range of biosensing applications, as well displaying an enhanced nucleobase recognition behaviour.[3,4]

XNA's that can be used to store and translate genetic information are of high interest in the field of biochemistry. Therose nucleic acid (**TNA**) a tetrose based replacement of the ribose sugar which is capable of self-hybridisation with itself can also cross-hybridise with DNA and RNA.[5,6] Futhermore, it can be polymerised by a natural DNA polymerase

(therminator) to create XNA oligonucleotides upto 80 nucleobases in length.[7] However, many XNA's are poor substrates for natural polymerases and reverse transciptases and as such much work has been invested in to the development of engineered polymerases and reverse transcriptase enzymes that can synthesise XNA polymers. Hexose nucleic acid (**HNA**), where the ribose sugar is replaced by a 6-membered sugar unit, has not only shown the ability to maintain and participate in sequence specific hydrogen bonding motifs with natural DNA, RNA and itself,[8] but can also be enzymatically polymerized using DNA templates and engineered polymerases (Figure 3.1a) as well as store genetic information that can be reverse transcribed into DNA using engineered reverse transcriptase enzymes (Figure 3.1b).[9,10] The ability to translate DNA information in to XNA polymers and vice-versa has been described by Pinheiro *et al* with the following XNA's:[10] therose nucleic acids (**TNA**); arabinose (**ANA**) and fluoro-arabinose (**FANA**) derivatives of DNA; cyclohexenyl nucleic acids (**CeNA**),[11,12] where the a 6-member sugar unit incorporates a double bond between the 5' and 6' position on the sugar unit and **LNA**.[10]

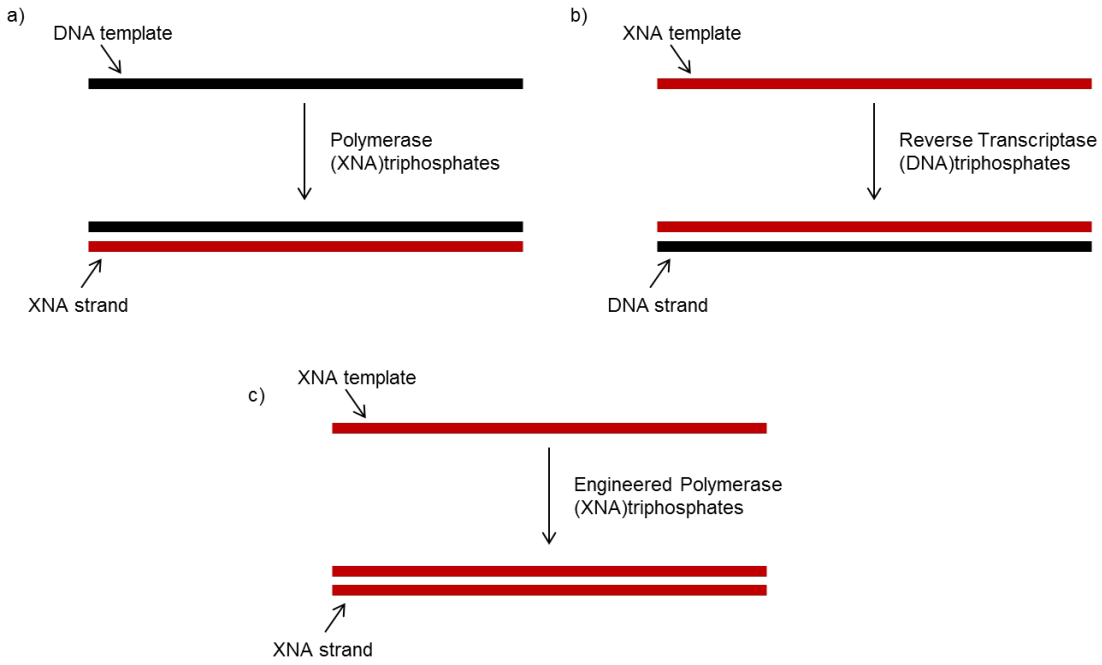


Figure 3.1 - A schematic representation of XNA's being able to participate in a) DNA templated polymerisation, b) reverse transcription. c) XNA to XNA replication.

Also of interest is the ability to template XNA synthesis using XNA templates (Figure 3.1c). Recently polymerases that can synthesise XNA strands in this way have been used for **FANA**, **HNA** and **CeNA** motifs.[10]

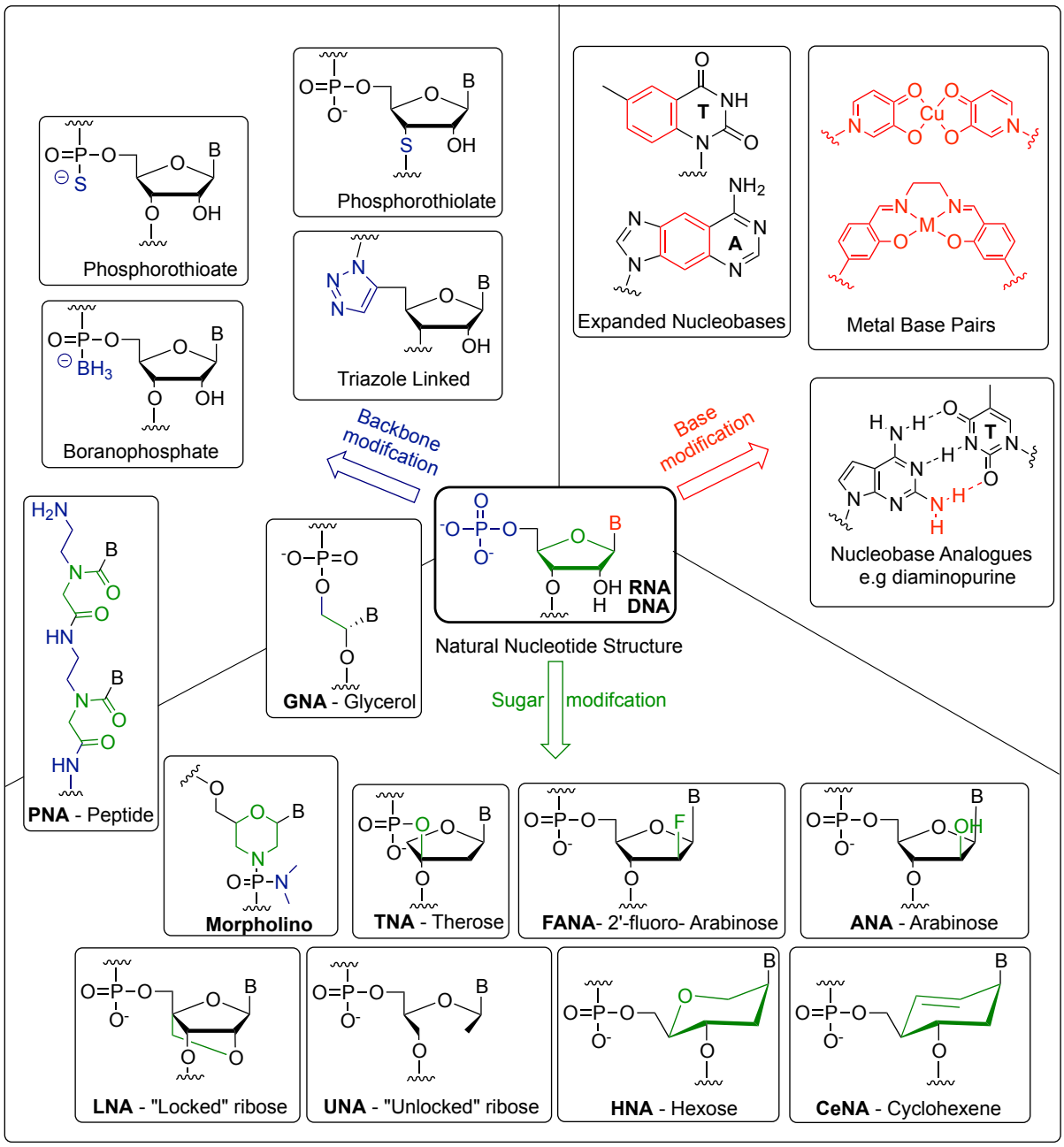


Figure 3.2 - Examples of developed XNA's where modifications to the sugar, nucleobase, backbone or a combination of these have been made. B = Nucleobase. References not mentioned in the text. ***GNA***,[13] ***Morpholino***,[14] ***UNA***,[15] *boranophospates,*[16] *expanded nucleobases.*[17,18]

Modifications to the phosphate backbone of DNA have also resulted in enhanced stability to enzyme degradation. In particular sulfur modifications to the non-bridging (phosphorothioates)[19] and bridging oxygens (phosphorothiolates)[20-22] of the DNA and RNA backbones notoriously provide stability improvements in this respect, but also provide the ability to append oligonucleotide strands to gold surfaces *via* Au-S chemistries.[23-25] More recently, Tom Brown and co-workers have developed triazole linked DNA oligonucleotides which can be conjoined by virtue of Cu-catalysed alkyne-azide coupling (so called 'click' chemistry) to form conjugates that can be transcribed into RNA. This approach has also made significant advancements in DNA synthesis technologies for producing synthesised oligonucleotides as long as genes or PCR amplimers.[26] Furthermore, this concept provides potential for the future design of unique DNA structures that were otherwise unobtainable using traditional DNA synthesis methods.[27]

Arguably, the most successful and most recognized XNA is peptide nucleic acid (**PNA**). First described by Nielsen *et al.* in 1991,[28] PNA represents a complete replacement of the natural sugar-phosphate-sugar motif of DNA with that of a simple repeating peptide unit, which bears a nucleobase appended to its side-chain.[29] Applications for PNA are vast due to its stability and lack of charge. Whilst this latter property makes it a challenge to transport PNA into cells, it has been utilized in many electrochemical[30-32] and optical[33] biosensing platforms due to both its stability and its significantly enhanced binding ability to DNA over natural nucleic acids (T_m - **PNA**•DNA > RNA•DNA > DNA•DNA).[29,34]

With genetic information being stored within the four canonical nucleobases; adenine, guanine, cytosine, thymine, the choices for information storage, building structure and controlling function could be considered quite limited when compared with amino acids of which there are twenty natural varieties. One derivative 2,6-diaminopurine (shown in Figure 3.2) has an additional amino functionality that allows it to form an extra hydrogen bonding interaction with thymine, creating more stable duplexes. In fact, recently diaminopurine has

been established as an extra-terrestrial nucleobase found on meteorites,[35] as well as being found in S-2L cyanophages which attack blue-green algae.[36,37] Over the past 20 years the groups of Carell and Shionoya have led the development of a whole series of metal binding ligands[38] that replace natural nucleobases, binding an array of metals such as copper[39], manganese[40] and silver[41]. Inspiration for their designs stemmed from the well known thymine-mercury-thymine interactions that is one primary reason for the mercury toxicity in nature.[42-44] A more detailed discussion of metal binding by nucleobases is included in Chapter 4.

3.1.2 Ferrocene

Ferrocene, an organometallic sandwich complex consisting of two cyclopentadienyl (Cp) rings encapsulating a single iron atom, was first synthesized in 1951 by Pauson and Kealy.[45] However its sandwich structure was not deduced until the following year by Robert Woodward and Geoffrey Wilkinson[46] and at the same time, independently by Ernst Otto Fischer.[47] For their contributions to the field metallocenes and the wider field of organometallic chemistry, Wilkson and Fischer were awarded the Nobel Prize in 1973.

Unlike most organic compounds, ferrocene is able to undergo a reversible one-electron oxidation, forming the stable ferrocenium cation (Figure 3.3). This property has led to the development of many electrochemical methodologies involving the use of ferrocene and its derivatives in a range of applications.

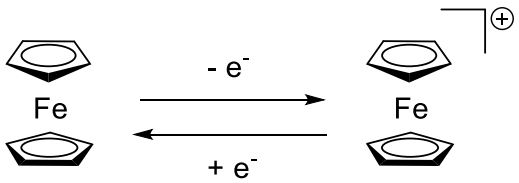


Figure 3.3 – The ferrocene redox couple.

Indeed, ferrocene is used by diabetics across the globe to probe blood glucose levels. In this assay, which was designed in the 1980's, a naturally occuring enzyme glucose oxidase oxidises glucose, which results in the release of two electrons to the enzyme. The enzyme passes these electrons to two ferrocenium cations which are generated at an electrode surface, and subsequently reduces ferrocenium to ferrocene (Figure 3.4). The amount of current required to convert ferrocene back to the ferrocenium is measured and used to determine the levels of glucose in the body.[48]

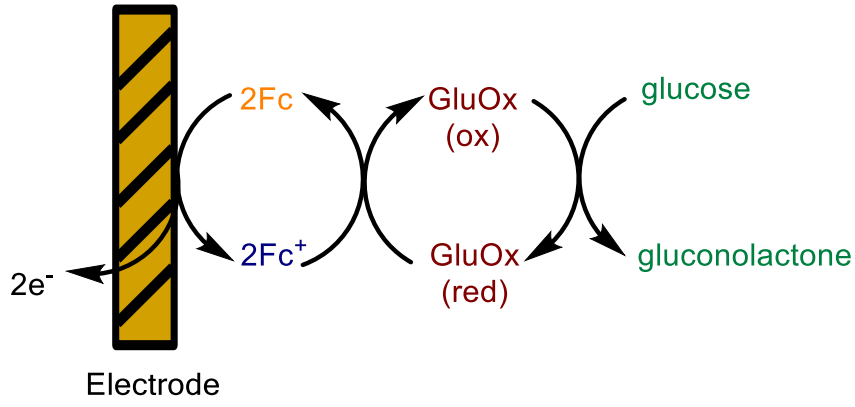


Figure 3.4 - A schematic representation of the redox processes involved in the ferrocene-based glucose sensor developed by Hill and co-workers.[48]

Ferrocene has also attracted considerable interest in the fields of organic synthesis and catalysis. Derivitisation of the Cp rings can be used to introduce two modes of chirality that have been utilised in the design of drugs,[49,50] catalysts[51] and chiral sensors.[52,53] Planar chirality is brought about by introducing two different groups onto one Cp ring (Figure 3.5). This relatively uncommon form of stereochemistry as well as the relatively low expense of synthesising ferrocene derivatives, has resulted in ferrocene becoming one of the most greatly investigated compounds in chemical research.

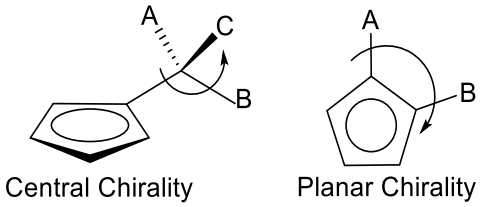


Figure 3.5 - Modes of chirality about derivatised ferrocene compounds.

3.1.3 Metallocene Incorporation into DNA.

Given its usefulness as an electrochemical tag, ferrocene has been used many times as a redox label for biomolecules.[54] In particular the labelling of oligonucleotides can be achieved *via* three main routes (described in more detail in Chapter 1.4):

- Chemical incorporation of a phosphoramidite *via* automated DNA synthesis.
- Enzymatic incorporation of a phosphate or triphosphate by DNA polymerases.
- Post DNA synthesis labeling *via* conjugation techniques such as NHS esterification or 'click chemistry'.

Examples of compounds used in these three approaches are shown in Figure 3.6. Two detailed reviews on the labelling of DNA with ferrocene have been published by Hianik[55] and Tucker.[56]

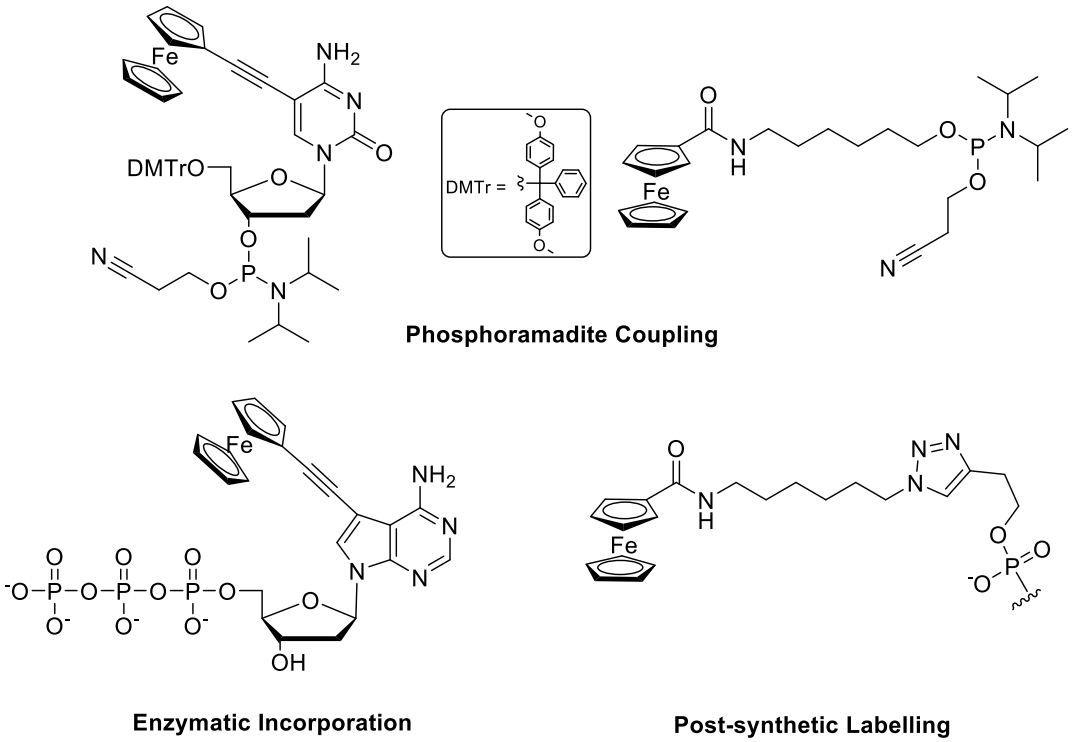


Figure 3.6 - Examples of ferrocene based oligonucleotide labels, showing three methods for incorporation.

Ferrocene tagged oligonucleotides were utilised in many of the early DNA based electrochemical sensors. Generally these sensors work by being sensitive to changes in DNA conformation as a result of binding a target of interest, be it a target strand of DNA,[24,57-59] or biologically relevant analytes using aptamers.[60,61] The first example, published by Ihara *et al*. described the use of a ferrocene (top right Figure 3.6) labelled strand that could electrochemically discriminate between duplex and triplex DNA structures.[62] This example set a great precedent for the development of a whole array of ferrocene DNA based sensors that would emerge over the next decade.[25,57,58,63-67] Whilst labelling with ferrocene was paramount to the development of these assays, the true potential of such designs was aided by the ability to immobilise such sensors onto electrode surfaces. Designs that relied on solution based investigations were thwarted by working concentrations and detection limits being well above those required for the detection of biologically relevant analytes. The group

of Kevin Plaxco (UC-Santa Barbara) which is at the forefront of the DNA sensor field, described one of the most well established and straightforward methods for DNA detection that uses a hairpin, molecular beacon approach. The design involved a thiol labelled oligonucleotide which forms a hairpin, with the 5' end of the oligonucleotide post synthetically modified with a ferrocene group (Figure 3.7). Upon hybridisation of target DNA, the hairpin opens up and binds its complementary strand, this moves the ferrocene tag away from the electrode, which subsequently reduces the rate of electron transfer rate and hence the signal diminishes; this so called 'Signal OFF' sensor has a detection limit of 5 fmol.[57]

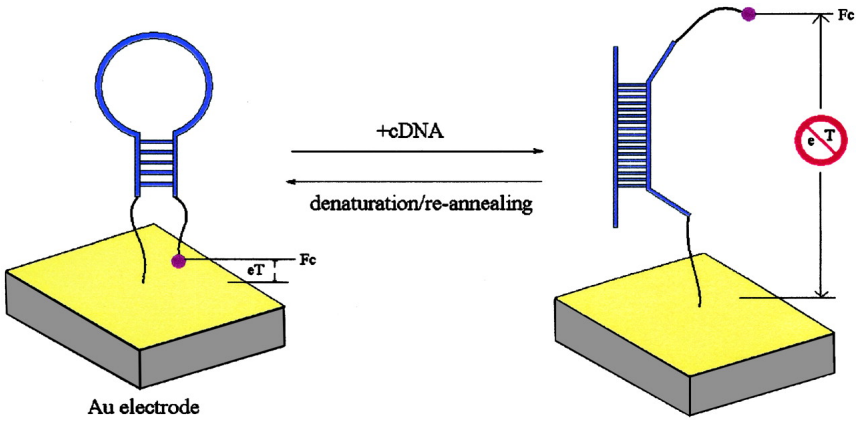


Figure 3.7 - Ferrocene labelled oligonucleotides for DNA detection by changes in conformation, published by Plaxco et al.[57]

In 2005, Inouye and co-workers developed end tagged ferrocene labels further by labelling the 5' end of an oligonucleotide with a ferrocene connected to both an isoquiline group (for stacking) and a dimidopyridine group (for hydrogen bonding). They were able to discriminate between subtle changes in DNA sequence (mutation, insertion and deletion of nucleobases), by virtue of small changes in duplex conformation and hence electron transfer rates from the ferrocene tag (Figure 3.8).[68-71]

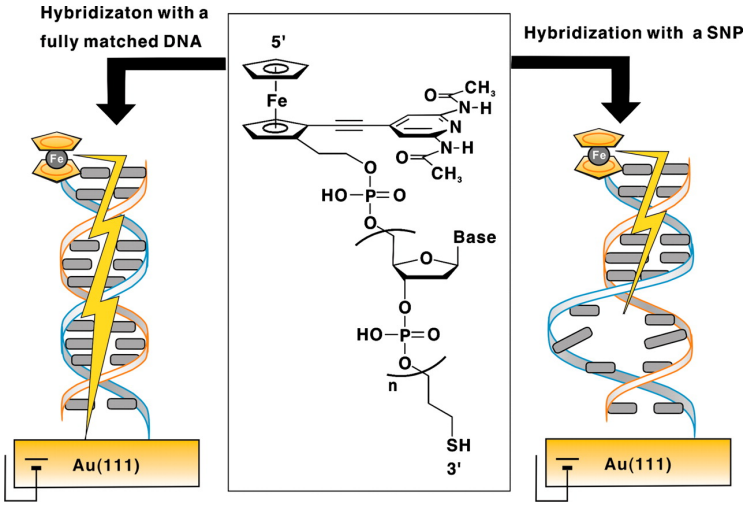


Figure 3.8 – A schematic of a electrochemical single nucleotide polymorphism (SNP) detection assay using a ferrocene-isoquinoline derivative tagged on DNA and immobilised on a gold electrode surface.[68]

The majority of ferrocene containing oligonucleotides that have been reported involve similar end tag (5', 3') labelling of strands. Ferrocene tags that lie in the middle of an oligonucleotide sequence by connecting two strands have to date been rare. In 2004 Chaix *et al.* reported the synthesis of a bis-substituted ferrocene phosphoramidite that could be incorporated directly into an oligonucleotide sequence in this way.[72,73] The unit consisted of a 3-carbon aliphatic linker connecting each Cp ring to the backbone of DNA. The incorporation of this motif in to the middle of a sequence was found to destabilise the duplex, with the level of destabilisation dependent on the position and number of incorporations.[74,75] The modified strands were shown to exhibit redox activity and respond directly to the addition of target strands of DNA by changes in current output (Figure 3.9).[72] In an effort to enhance electrochemical sensitivity, multiple additions (up to 8) have been incorporated,[74,76] and some systems have been immobilised on to an electrode surface.[23]

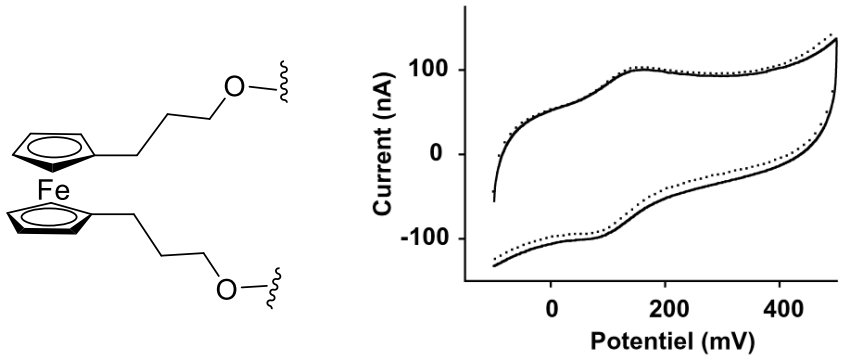


Figure 3.9 - Bis-substituted ferrocenylalykl motif incorporated into DNA via phosphoramidite chemistry and automated DNA synthesis. Solid line – single stranded, dashed line – double stranded.[72]

In an extension of this work, Ihara and co-workers reported the incorporation of the same C3 ferrocene moiety shown in *Figure* 3.9 into DNA, as well as two other extended versions consisting of polyethylene glycol units. Their work also showed that incorporation resulted in a destabilisation of the duplex. However, they also found that binding to a DNA target strand resulted in small negative shifts in the ferrocene centred oxidation potentials. This was ascribed to stabilisation of the ferrocenium cation by the negatively charged DNA backbones.[77]

3.1.4 Ferrocenyl Nucleobases

The labelling of nucleobases with ferrocene has become an emerging field of interest in the past decade. This has gained prominence through the desire to design new drugs for the treatments of cancers and viral ailments using nucleoside analogues and nucleobase derivatives. Nucleobase analogues bear a primary alcohol, nucleobase and a spacer group (Figure 3.10). Azidothymidine (AZT), (the most commonly prescribed nucleosidic drug in clinics today), is still commonly used for the treatment of HIV/AIDS. AZT works by blocking the revere transcription pathway that the HIV virus uses to replicate, having hijacked a patient's cell. Such has been the success of AZT and other nucleoside analogs such as

aciclovir and gemcitabine, that there is continued interest in the development of new nucleoside units that bear different side chains and functional groups.

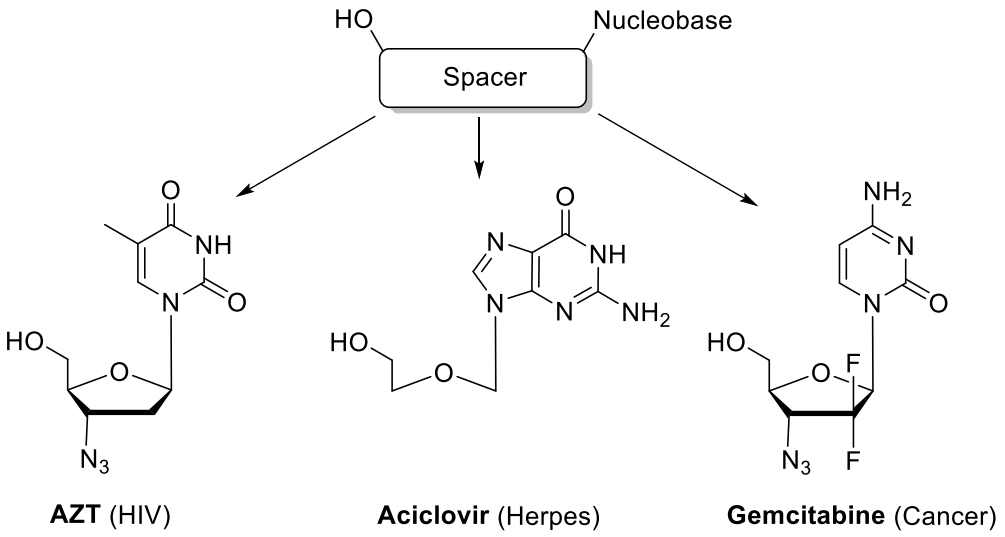


Figure 3.10 – Examples of nucleoside analgoues AZT, aciclovir and gemcitabine and the ailments they are used to treat.

Ferrocene itself is considered by some to have anti-proliferation properties caused by the release of radicals by the *in vivo* oxidation to Fe^{III} which causes oxidative damage of DNA, proteins and lipids.[50,78-80] Additionally the lipophilic nature of ferrocene and its derivatives makes it a useful group for faciliating the crossing of cell membranes, which remains a challenge for many drug compounds.

Figure 3.11 shows the metallocene compounds currently described in the bioorganometallic chemistry literature which bear single nucleobases; cytosine (**3.1**, **3.8**),[81,82] adenine (**3.2**, **3.9**),[81,83] uracil(**3.3**, **3.5** & **3.7**),[84,85] thymine (**3.4** & **3.10**),[83,84] and guanine (**3.4**)[86]. These compounds, albeit not strictly nucleoside mimics, showed some anticancer activity and interesting electrochemical properties.

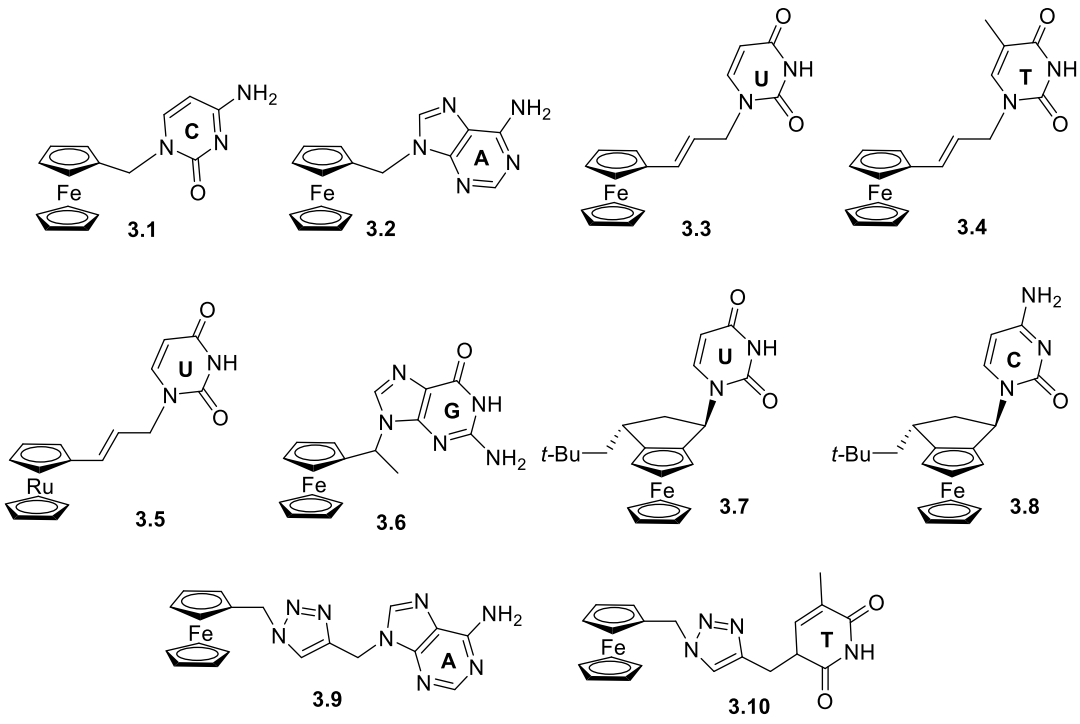


Figure 3.11 - Examples of mono-substituted ferrocene and ruthenocene derivitised with nucleobases. References detailed in text.

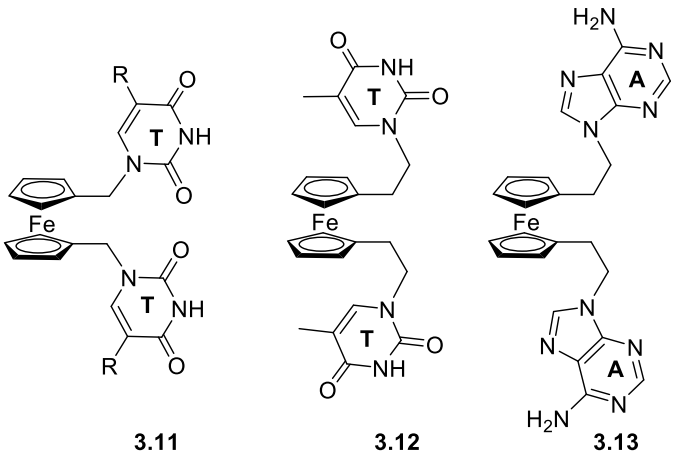


*Figure 3.12 - Bis-subtituted ferrocenyl nucleobases. Published by Ganesh[87] (**3.11**) and the Tucker Group (**3.12, 3.13**).[88]*

A step further has been the development of bis-substituted ferrocene-nucleobase derivatives bearing thymines (**3.11** and **3.12**) and adenines (**3.14**) shown in Figure 3.12. These motifs were designed mainly for the formation of supramolecular hydrogen bonded

networks rather than for their drug potential.[87,88] However, these designs have since formed the foundation of future designs of organometallic DNA mimics.

More recently, the Tucker Group have built on the work of ferrocene-nucleobase conjugates to synthesise a new range of ferrocene based nucleoside analogs (**3.18**)[49]. These motifs incorporate both a nucleobase (thymine/adenine) and a primary alcohol. From structure activity relationships against cancer cell lines (murine leukaemia - L1210, HeLa and human T-lymphocyte – CEM), it was found that both functionalities are required to produce good anticancer activity (Figure 3.13). The observed cytostatic activities are comparable to cis-platin which is the most heavily used chemotherapy drug in clinics today.[49]

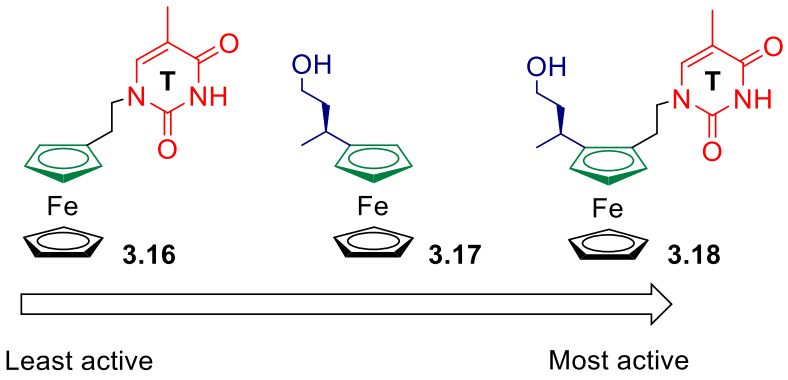


Figure 3.13 - Ferrocenyl nucleobase analogs for anticancer activity.[49]

3.1.5 Ferrocene Nucleic Acid (FcNA)

In 2012, the Tucker Group described the first example of an organometallic nucleic acid mimic.[89] The hypothesis for this mimic stemmed from the fact that the distance between adjacent Cp rings in the ferrocene structure (3.3 Å) is similar to that between stacked base-pairs in duplex DNA (3.4 Å). The tetra-substituted unit (shown in Figure 3.14), was designed to replace/mimic two nucleobase units (sugar-phosphate-sugar) of DNA and consists of a single ferrocene unit bearing two C3 linkages to primary alcohols and two C2 linkages to thymines. The four-fold substitution pattern meant that chirality had to be introduced in the

monomer units by way of an alpha methyl side chain appended to the C3 diol linkage; the central chirality at this position was used during the synthesis to direct the diastereoselective ortho substitution of the nucleobase linker to create the planar chirality. Furthermore, planar chirality is naturally imparted by the introduction of this functionality. This design was based on computational modelling studies carried out by Dr. John Wilkie at the University of Birmingham.[89]

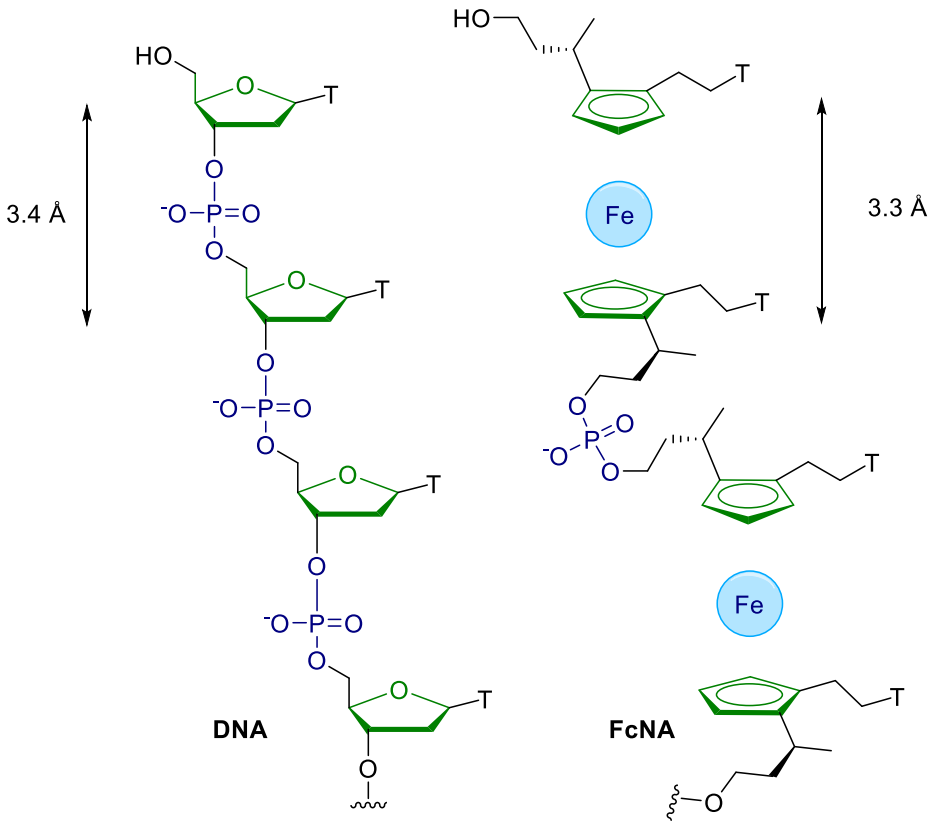


Figure 3.14 - Structures of DNA and FcNA showing how it mimics a dinucleotide.

Phosphitylation of one primary alcohol and tritylation of the other allowed for the use of the monomer in an automated DNA synthesiser. It was successfully oligomerised to create an octamer of the unit ((**FcTT$_{SS}$**)$_8$), which was subsequently purified by HPLC and characterised by mass spectrometry.[89] Although this particular oligomer was unable to form stable duplexes with its DNA counterpart (**A$_{16}$**), it did show *quasi*-reversible electrochemical behaviour (Figure 3.15) indicative of a multiple electron transfer process for each strand.[89,90]

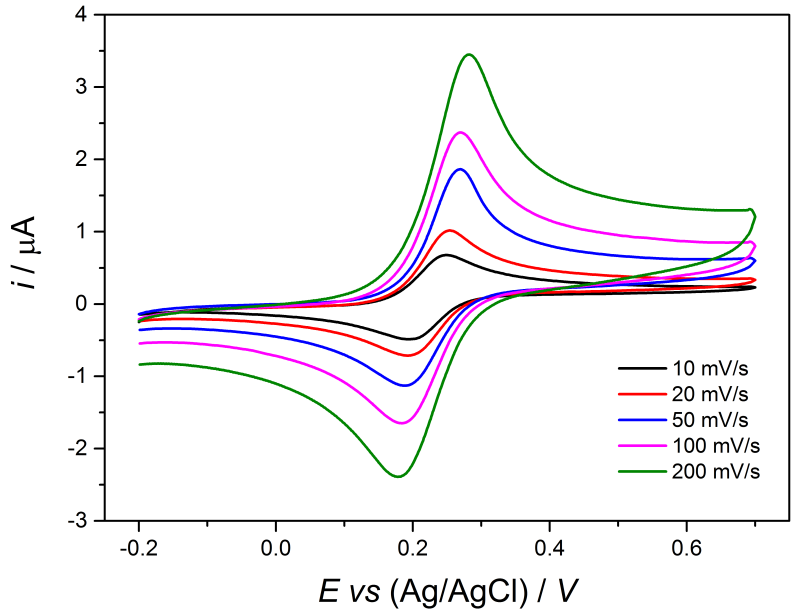


Figure 3.15 - Cyclic Voltammograms of ***(FcTT$_8$)$_{SS}$*** *an FcNA octamer.*[89]

3.1.6 Conclusions and Project Aims

The development of nucleic acid mimics and modifications to the major components of DNA has brought about the ability to control binding strength, discrimination and function through an expansion of the toolbox available to the biological chemist. Ferrocene, a redox active compound has shown great potential for its use in the field of bioorganometallic chemistry as an electrochemical tag and as a component in the design of new drugs. The Tucker group's work on the development of FcNA, showed the first example of a wholly organometallic structural mimic of a nucleic acid. Whilst specific interactions with DNA were not observed in the first example of an oligomeric FcNA system, the design principle and properties of FcNA open up a whole host of application possibilities for areas that include biosensing, DNA nanostructures and DNA nanowires.

The work discussed in this chapter was aimed at building upon the previous work within the group to develop, rationalise and optimise the FcNA system. Rather than using a wholly organometallic set of strands, this would be achieved by a single incorporation approach, in which just one FcNA unit is incorporated into the backbone of a DNA oligonucleotide so that

the flanking nucleobases can anchor the unit into a duplex environment (Figure 3.16). By controlled changes to the FcNA unit structure and its stereochemistry it was anticipated the duplex stabilisation and base discrimination behavior could be maximized. Furthermore, through analytical techniques such as circular dichroism (CD) spectroscopy, UV-vis spectrsoscopy and electrochemistry, the assessment of the structural effects of incorporating FcNA units into DNA could be probed in a systematic manner.

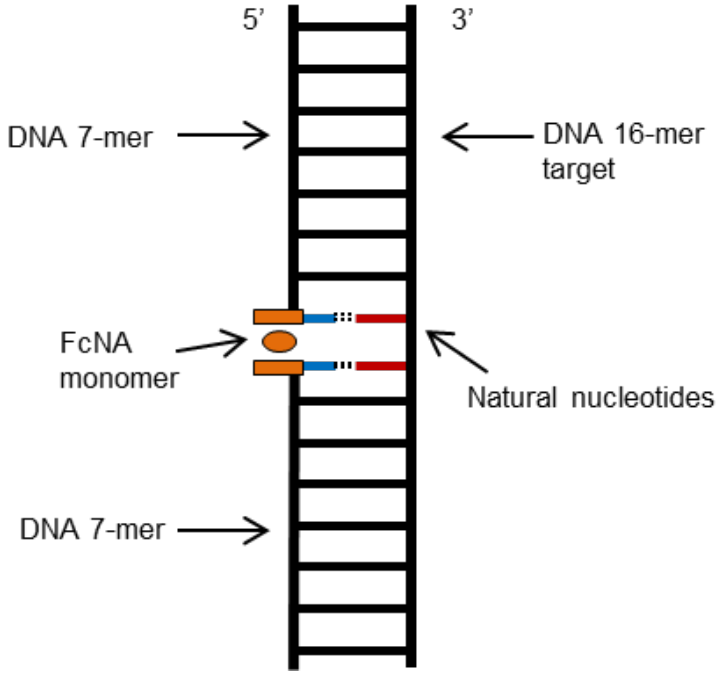


Figure 3.16 - Schematic representation of the single incorporation of FcNA monomers into a DNA backbone.

3.2 Results and Discusson

3.2.1 Design and Synthesis

All FcNA monomers were synthesised according to the strategy shown in Figure 3.17 by Dr. Huy Van Nguyen, Dr. Antoine Sallustrau or Jonathan Kedge and subsequently handed over at the phosphoramidite stage for incorporation into DNA *via* automated DNA synthesis. The synthesis of these monomers was generally achieved in less than 20 synthetic steps. Full details of the synthesis are beyond the scope of this discussion but information regarding their design and characterisation can be found in the cited references.[89,91] It is worth noting that since the initial publication of the monomer synthesis, the route has been further developed and optimised and will be published in due course.

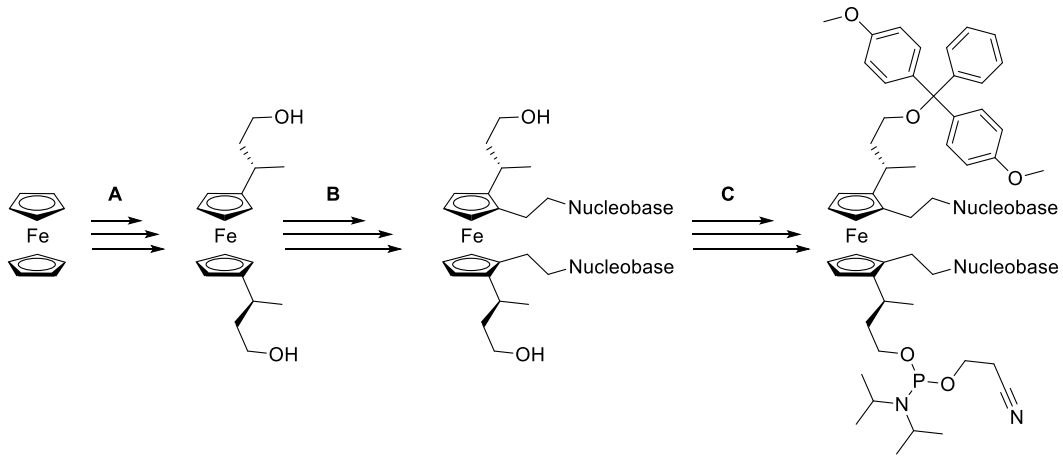


Figure 3.17 - Synthetic strategy for synthesis of FcNA monomers. ***A*** *- Appending linkages for connection to DNA backbone and introduction of planar chirality if required.* ***B*** *- Appending of nucleobases and introduction of planar chirality.* ***C*** *- protection with DMT and addition of reactive phosphoramidite functionality. Synthesis carried out by Dr. Huy Van Nguyen, Dr. Antoine Sallustrau or Jonathon Kedge.*

The initial FcNA design previously reported formed the foundation of the monomer design. In order to study the effects of each functional group and mode of chirality on the duplex stability and nucleobase recognition, a number of new monomeric units were

synthesised. It is important to understand the nomenclature associated with these units. Figure 3.18 depicts a schematic representation of the principle structural aspects of the FcNA monomer unit and the resulting nomenclature. It is important to note that whenever two or more substituents are appended to each Cp ring, planar chirality (**pYpY** = **pSpS**/**pRpR**) is naturally induced (please refer to Figure 3.5). The central chirality (***XX*** = ***RR/SS***) comes from the methyl group in the α position. As well as controlling chirality, the synthetic routes allow for control over the length of the aliphatic linkers to both the nucleobase and the alcohol synthetic functionalities, the lengths of which are defined as nX and cZ respectively. Finally the synthetic approach used for the addition of the nucleobase (**BB**) to the unit allows for control over the nucleobases that are appended. However at this stage, the synthetic route only allows for a homo-nucleobase strategy, where both appended bases are the same (e.g **TT**).

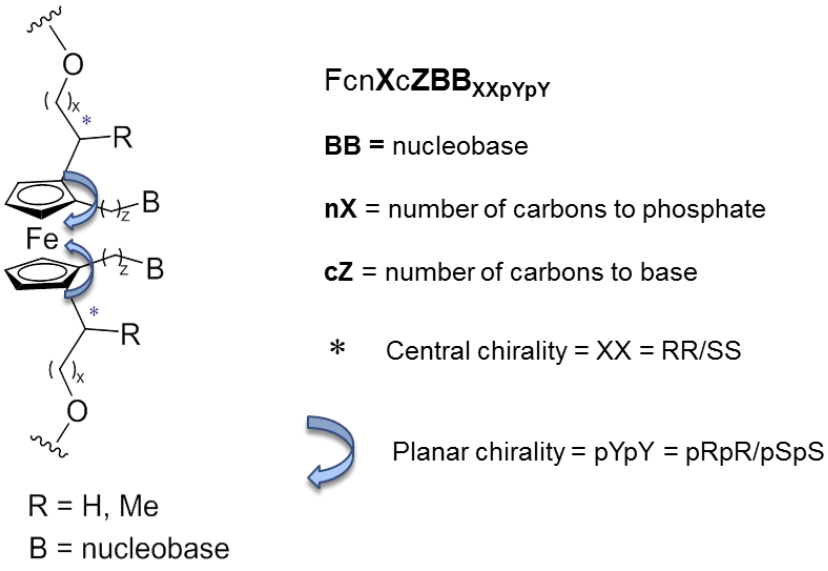


Figure 3.18 - Schematic representation of FcNA systems, showing the points of clarity and nomenclature.

Throughout the synthesis of these monomers, the chiral purity was checked by chiral-HPLC. An example of the CD spectra comparing two enatiomerically pure stereoisomers prior to the addition of the DMT and phosphoramidite groups is shown below (Figure 3.19).

The spectra are characterised by three bands at 225, 275 and 450 nm which are attributed to π-π* transitions for the Cp rings and the nucleobases, and the d-d transition of ferrocene respectively. The sign of the bands is dependent on the chirality of the monomer itself.

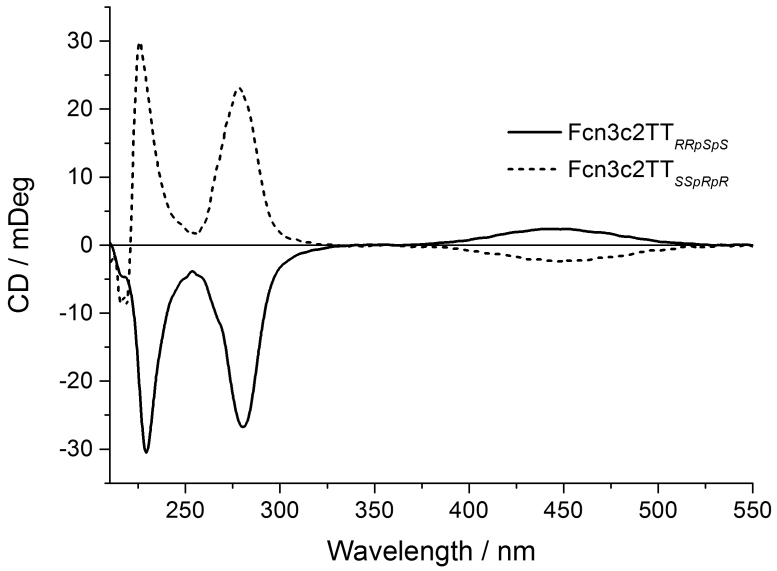


Figure 3.19 - CD spectra of the two stereoisomers of ***Fcn3c2TT$_{XXpYpY}$*** *at the diol stage. 400 μM in MeCN. Solid line –* ***Fcn3c2TT$_{RRpSpS}$****. Dashed line –* ***Fcn3c2TT$_{SSpRpR}$****.*

As detailed previously, in order to establish the most suitable structure and functionality for the FcNA units in binding natural DNA, the monomeric units were incorporated into the central part of a 15-mer oligonucleotide which has been studied extensively by the Tucker Group.[92-94] The sequence, **S1**, (shown in Table 3.1) consists of a ferrocene unit flanked by two sequences of 7 bases. Together these have equal numbers of A-T and G-C base pairs. Previous studies have shown there to be no self complementary behaviour when they are joined by a T base to form a 15-mer. Its full complementary strand, **S2**, was used to investigate the binding affinities where two nucleobase are inserted opposite the ferrocene unit to give a 16-mer. Throughout the investigations, the flanking nucleobases remain constant, and only the two central nucleobases, **XX** and **YY** in **S1** and **S2** respectively are altered.

*Table 3.1 – Sequences **S1** and **S2** where XX = FcNA/AA/TT/CC/GG/AbAb and YY = FcNA/AA/TT/CC/GG/AbAb/UU*

Strand	Sequence (5'-3')
S1XX	TGGACTC-**XX**-CTCAATG
S2YY	CATTGAG-**YY**-GAGTCCA

All FcNA monomers synthesised were successfully incorporated into the backbone of DNA, purified using reversed-phase HPLC (RP-HPLC) and characterized by mass spectrometry (see the experimental Chapter 7.2.2 for more details). Ferrocene labelled oliognucleotides are readily separated from truncated/failed sequences because of the enhanced lipophilicity afforded by the monomer unit. Characterisation of all the oligonucleotides used in this chapter can be found in appendix 1.

3.2.2 Storage and Stability

All synthesised oligonucleotides were stored at -20 °C in ultra-pure MilliQ water and separated into small aliquots to ensure the least amount of exposure to heat and thaw cycles. It was noticable that these strands were susceptible to oxidation unless stored in this manner; a further discussion of this phenomenon is discussed in Chapter 3.2.4.

3.2.3 Optimising the Structure of Thymine modified FcNA

Thermal melting profiles of DNA are used regularly as a technique to evaluate the overall stability of a DNA structure (for a detailed discussion of this methodology, see Chapter 2). By comparing the relative T_m values obtained when varying the FcNA unit and its opposing nucleobases, it is possible to gain an insight into the contributions each unit has on stabilising the duplex, either by base stacking contributions or through hydrogen bonding interactions.

3.2.3.1 FcNA without nucleobases – an abasic control

A ferrocene bearing no nucleobases, was synthesised and is described throughout as **FcHH*RR/SS*** (shown in Figure 3.20). It still contains a 3-carbon linker to the primary alcohol as well as a methyl in the α position. However as it lacks nucleobases it only has central chirality and has no planar chirality.

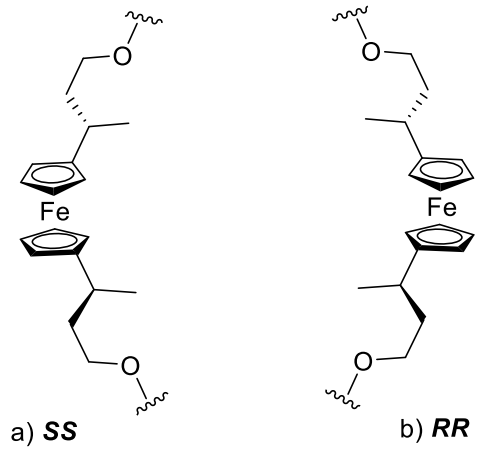


Figure 3.20 – The two stereoisomers of control FcNA unit, ***FcHH a) SS and b) RR.***

The **S1** strands containing the two stereoisomers of **FcHH** (***RR*** and ***SS***) were hybridised with different **S2** target strands. Table 3.2 shows the resulting T_m values recorded using UV-vis spectroscopy. As expected, the melting temperatures change by only small amounts, with any small fluctuations possibly caused by a change in nucleobase size opposite the ferrocene unit. Notably the introduction of the ferrocene unit is not significantly destabilising to the overall structure when compared with the DNA equivalent (**S1AbAb**) which has a double abasic site (Figure 3.21), and in some cases the overall stability is greater. In addition the central chirality of the two stereoisomers has little or no bearing on the overall stability of the duplexes. As expected for the abasic **S1FcHH** units only small differences were observed.

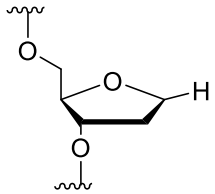


Figure 3.21 - The abasic nucleotide monomer incorporated into S1AbAb.

*Table 3.2 - T_m values for **S1XX•S2YY** when **XX** = **FcHH$_{RR/SS}$** and **YY** = AA, AG, GG, TT. DNA control **S1AbAb** is also shown. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C **S2YY=**			
Strand	**S2AA**	**S2AG**	**S2GG**	**S2TT**
S1FcHH$_{SS}$	37.5 (±0.2)	37.0 (±0.2)	36.5 (±0.3)	36.0 (±0.2)
S1FcHH$_{RR}$	37.5 (±0.2)	37.5 (±0.3)	36.5 (±0.3)	36.5 (±0.0)
S1AbAb	37.0 (±0.0)	37.0 (±0.0)	37.0 (±0.0)	34.0 (±0.0)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

To support these studies, molecular dynamic (MD) simulations (min. 5 ns) were carried out by Dr. Jean-Louis Duprey on the two **S1FcHH** systems (with **S2AA**) and representations are shown below in Figure 3.22. Interestingly, and despite the similarities in T_m values, they show quite different structures. The ***SS*** isomer appears to penetrate more deeply into the cavity created by the absence of nucleobases, which causes the Cp rings to sit orthogonal to the base stack. As a consequence, the methyl groups also sit inside the hydrophobic core of the duplex. In contrast, the ***RR*** isomer would appear to protude more from the duplex core, and follow the helical twist of DNA more naturally; the methyl groups again are situated in the hydrophobic cavity.

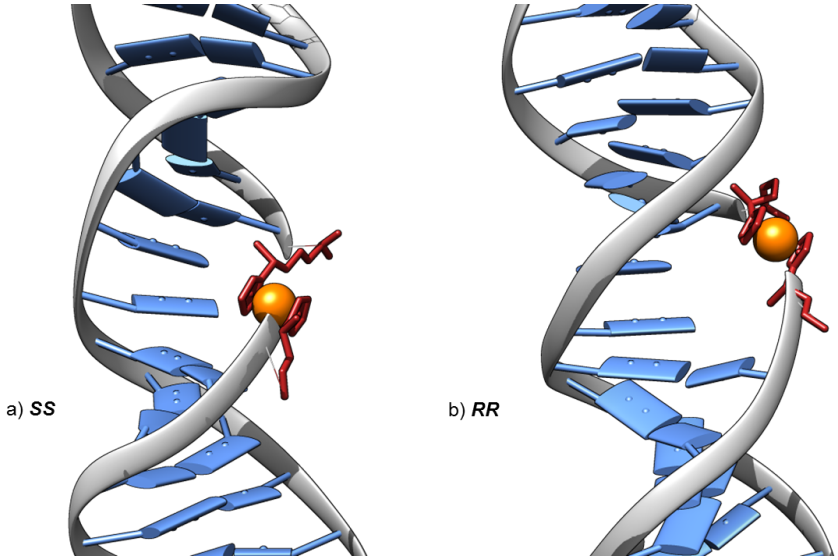


*Figure 3.22 - Molecular Dynamic simulations (5 ns) of **S1Fcn3HH** a) **SS** and b) **RR** duplexed with **S2AA** - models created by Dr. Jean-Louis Duprey using Insight II. Simulations portrayed using Chimera.*[95]

3.2.3.2 *FcNA containing two carbons to thymine and three carbons to phosphate*

Analagous systems containing T bases were investigated including the published monomeric unit **S1Fcn3c2TT$_{SSpRpR}$** (Figure 3.23a).[89] The resulting T_m values (Table 3.3) showed a significant advance on that of the abasic control **S1FcHH$_{SS}$**, with the introduction of the two thymines resulting in an increase in the T_m of 3.5 °C for the duplex with the fully complementary strand **S2AA**.

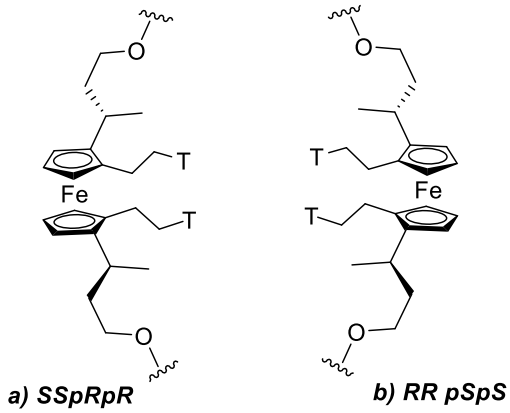


*Figure 3.23 - The structure of FcNA monomer **Fcn3c2TT$_{XXpYpY}$** incorporated into **S1**, showing the two different stereoisomers a) **SSpRpR** and b) **RRpSpS**.*

A greater increase was observed, for the ***RR*** isomer, **S1Fcn3c2TT$_{RRpSpS}$** (Figure 3.23b) with an increase of 5.5 °C over its abasic equivalent **S1FcHH$_{RR}$.** The 2 °C increase over the ***SS*** isomer suggests that planar chirality plays an important role in the overall stability of the duplexes. Whilst the increases represented a significant enhancement, it was important to investigate whether base pairing across the duplex was apparent by varying the bases opposite. This would include different bases of the same size (e.g. GG) and smaller bases (e.g. TT).

*Table 3.3 - T_m values for **S1XX•S2YY** when **XX** = **FcHH$_{RR/SS}$** and **YY** = AA, AG, GG, TT. DNA control **S1TT** is also shown. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C **S2YY=**			
Strand	**S2AA**	**S2AG**	**S2GG**	**S2TT**
S1Fcn3c2TT$_{SSpRpR}$	41.0 (±0.3)	39.5 (±0.3)	39.5 (±0.3)	39.5 (±0.3)
S1Fcn3c2TT$_{RRpSpS}$	43.0 (±0.1)	42.0 (±0.0)	41.0 (±0.0)	40.5 (±0.2)
S1TT	56.0 (±0.2)	48.5 (±0.3)	44.0 (±0.0)	41.5 (±0.2)
S1AbAb	37.0 (±0.0)	37.0 (±0.0)	37.0 (±0.0)	34.0 (±0.0)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

Whatever the system studied, it is just worth noting that a large destabilisation is found when compared with the fully complementary duplex of unmodified DNA (**S1TT•S2AA**). This is not entirely unexpected, given that the ferrocene unit is more ridged than natural DNA, and which might cause the flanking nucleobases to buckle or loosen their base stack to accommodate the new motif. The ***RR*** isomer shows the most positive signs of a H-bonding interaction, having the highest T_m value of 43 °C and the largest difference (2 °C) between complementary (**S2AA**) and double mismatch sequences (**S2GG**) that also contains a purine that would stack effectively. Significant changes in the size and hence stacking ability of the opposing nucleobases has a clear effect on the overall T_m given that **S2AA** vs. **S2TT** results in a 2.5 °C reduction for the ***RR*** isomer. The T_m value of this unit with the single complementary base pair target **S2AG** shows a melting temperature of 42°C which suggests that one A-T base pair has been lost.

MD simulations comparing the duplexes of both stereoisomers show two differing structures that support the observed differences in T_m values. As with the abasic derivatives **S1FcHH**, the ***SS*** shows a kink at the ferrocene unit in the line of the backbone presumably to maximise the stacking interactions with the two complementary nucleobases. As a result the models suggest an un-twisting and distortion of the flanking bases in the sequence, which results in a lengthening of the duplex. In comparison, the ***RR*** isomer seems to follow the natural twist of the DNA backbone more closely with less strain imparted on the structure as a result (Figure 3.25). It appears that the opposing nucleobases do not have to alter their position as much to accommodate the unit, and therefore the duplex is tighter and hence shorter in length.

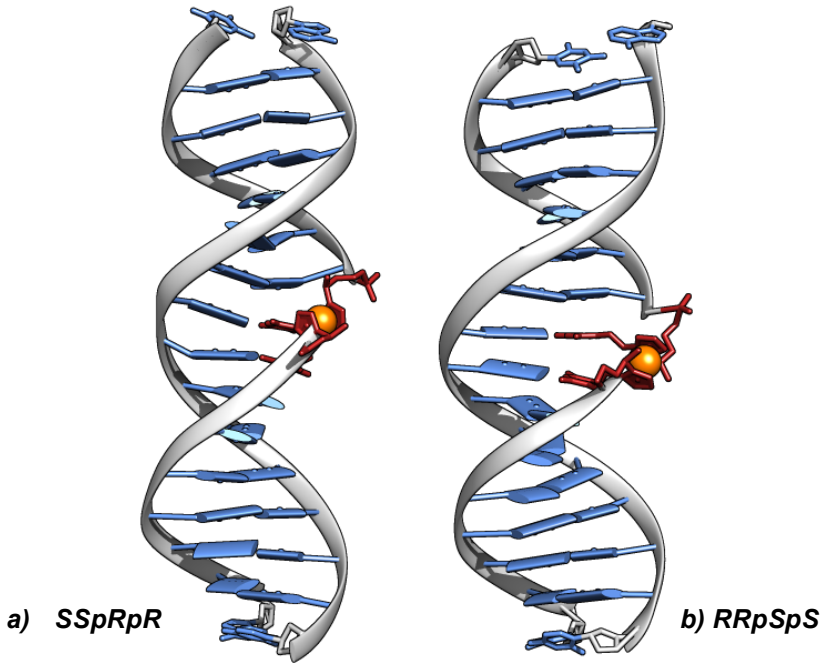


*Figure 3.24 - Molecular dynamic simulations (5 ns) of **S1Fcn3c2TT** with **S2AA** for both stereoisomers (a) **SSpRpR** (b) **RRpSpS** depicting the difference in duplex size. Visualised using Chimera II.*[95]

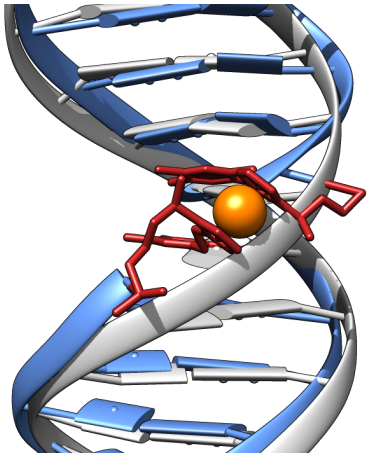


Figure 3.25 - MD simulation (5 ns) showing the effect of the incorporation of ***Fcn3c2TT$_{RRpSpS}$*** *(red) into* ***S1*** *(light blue) when compared with* ***S1TT*** *(grey) and duplexed with* ***S2AA****. Visualised using Chimera II.*[95]

3.2.3.3 Increasing the linker length to the nucleobase

By increasing the number of carbons in the aliphatic linker between the Cp rings and nucleobase, it was hoped that a correlation between duplex stability and length would begin to arise. In particular it was anticipated that an additional carbon would increase the reach and enhance the fit of the nucleobase stack. The monomers **Fcn3c3TT** (***SS*** & ***RR***) (Figure 3.26) were therefore incorporated into strand **S1** and the resulting T_m values are shown in Table 3.4.

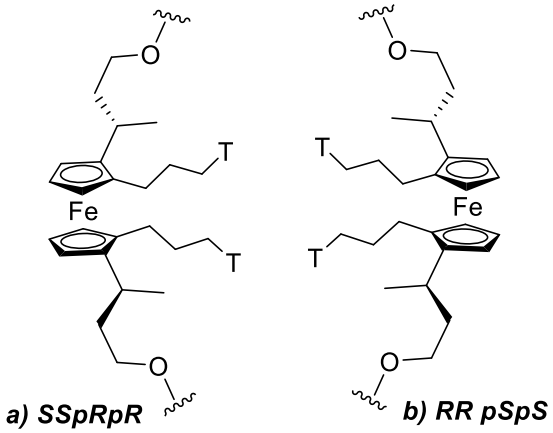


Figure 3.26 - The structure of FcNA monomer ***Fcn3c3TT*** *showing the 3-carbon linker to the bases and the two different stereoisomers (a)****SSpRpR*** *and (b)****RRpSpS****.*

*Table 3.4 - T_m values for **S1Fcn3c3TT** for both stereoisomers when duplexed with **S2YY**. Comparable systems with **S1TT** and **S1Fcn3c2TT$_{RRpSps}$** are also shown. 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C S2YY=			
Strand	**S2AA**	**S2AG**	**S2GG**	**S2TT**
S1Fcn3c3TT$_{SSpRpR}$[#]	43.0 (±0.1)	41.5 (±0.6)	40.5 (±0.2)	40.5 (±0.2)
S1Fcn3c3TT$_{RRpSpS}$[#]	41.5 (±0.2)	41.0 (±0.3)	40.5 (±0.2)	40.5 (±0.4)
S1TT	56.0 (±0.2)	44.5 (±0.3)	44.0 (±0.0)	41.5 (±0.2)
S1Fcn3c2TT$_{RRpSpS}$	43.0 (±0.1)	42.0 (±0.0)	41.0 (±0.0)	41.5 (±0.2)

* Highlighted values represent 'fully complementary' sequences. # T_m value increased with each ramp suggesting an instability of the oligonucleotide to the conditions. The temperature for the first ramp is therefore used. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

The initial observations of the T_m values suggested that the 3-carbon linkage gave slightly enhanced T_m values compared with those of the C2 moiety, albeit for the opposite isomer (***SS***). However upon closer inspection of the individual ramps of the T_m experiments, it became apparent that these oligonucleotides were unstable to the rigours of the T_m experiments, with the values increasing (e.g. from 43-46 °C) with each simultaneous denaturing and annealing ramp (typically the denaturing ramp is always slightly lower due to hysteresis).[96] This gave the first indication that FcNA systems may not be able to withstand completely the aqueous and salt conditions required for studying nucleic acid based systems. One possible explanation is that its oxidation is made easier by the increased +I inductive effective of the longer alkyl linker. Due to this effect, and in order to compare with the other systems, the T_m value for the first denaturing ramp was used (Table 3.4). MD simulations of the two stereoisomers suggest markedly different structures. The ***SS*** stereoisomer appears to organise the two thymines in line with both the helical twist and staircase nature of the nucleobase stack, as shown in Figure 3.28. Conversely, for the ***RR*** stereoisomer, the thymines are buckled which results in a skew in the parallel stack, which would be consistent with supporting the lower duplex stability observed for this system.

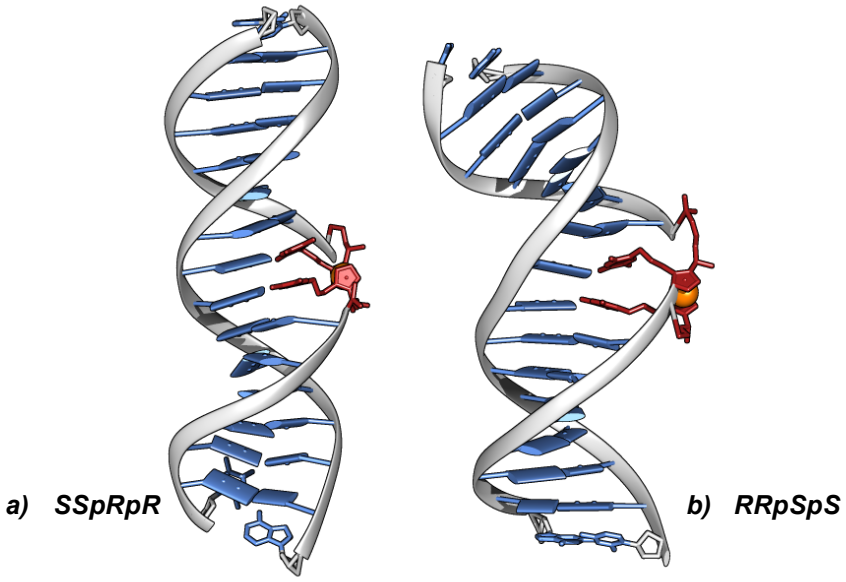


Figure 3.27 - MD simulations (6 ns) of the two isomers of ***Fcn3c3T$_{XXpYpY}$*** *when incorporated into* ***S1*** *and duplexed with* ***S2AA****. A)* ***SSpRpR*** *b)* ***RRpSpS***

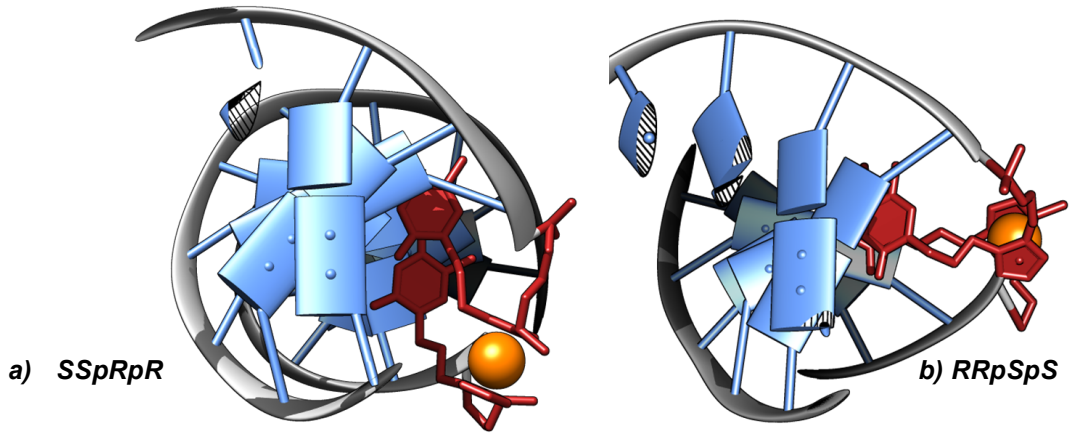


Figure 3.28 - MD simulations (6 ns) of the two isomers of ***Fcn3c3T$_{XXpYpY}$*** *when incorporated into* ***S1*** *and duplexed with* ***S2AA****, showing a cross-section of the duplex highlighting the difference in nucleobase stack of the FcNA monomer and flanking nucleobases. A)* ***SSpRpR*** *b)* ***RRpSpS***

3.2.3.4 Removing the methyl and central chirality

Whilst studying the length of the nucleobase linker was important, it was also of interest to study the effect of removing the central chirality and more specifically, the hydrophobic methyl groups, which appeared to be twisting into the duplex core. The loss of the methyl means that the priority for naming the planar chirality changes. ***pRpR*** derivatives mimic the

helical twist of the previously described ***pSpS*** derivatives and vice versa (Figure 3.29). Once again, these derivatives were successfully incorporated into S1 *via* automated DNA synthesis.

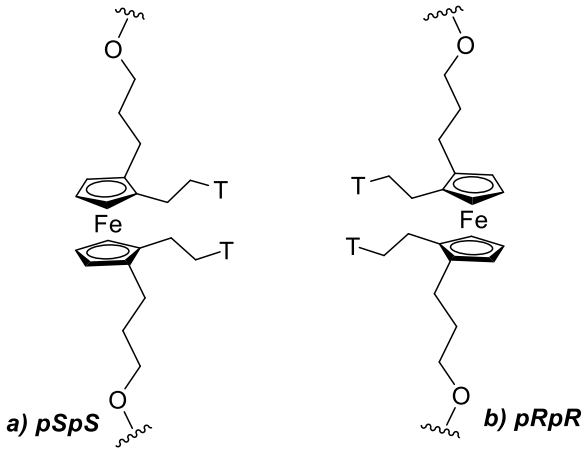


Figure 3.29 - The structure of FcNA monomer ***Fcn3c2TT$_{pYpY}$*** *showing both stereoisomers (a)* ***pSpS*** *& b)* ***pRpR****) synthesised.*

The T_m values obtained (Table 3.5) suggest that the combination of methyl and central chirality does make a contribution to the overall stability of the duplex. Removing the methyl group has a detrimental effect on the duplex stability, with the fully complementary system (**S1Fcn3c2TT$_{pRpR}$•S2AA**) reducing its T_m by 1.5 °C compared to with **S1Fcn3c2TT$_{RRpSpS}$•S2AA**. This gives further support to the idea that the methyl group contributes to the stability of the system. Furthermore, the base discrimination (**S2AA** vs **S2GG**) is almost negligible for both stereoisomers. For this unit it appears that the size of the base opposite does not affect duplex stability greatly, especially for the ***pSpS*** isomer. The MD models (Figure 3.30) suggest that the unit imparts a slight untwisting of the duplex so that the cavity afforded by the incorporation is too large for stacking contributions to stabilise the duplex in this region (i.e. GG should stack better than TT).

*Table 3.5 – T_m values for **Fcn3c2TTpYpY** in **S1** when duplexed with **S2YY**. 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C S2YY =			
Strand	**S2AA**	**S2AG**	**S2GG**	**S2TT**
S1Fcn3c2TT*pSpS*	40.5 (±0.2)	39.5 (±0.3)	40.0 (±0.4)	40.5 (±0.2)
S1Fcn3c2TT*pRpR*	41.5 (±0.0)	41.5 (±0.3)	41.0 (±0.0)	40.5 (±0.3)
S1TT	56.0 (±0.2)	44.5 (±0.3)	44.0 (±0.0)	41.5 (±0.2)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

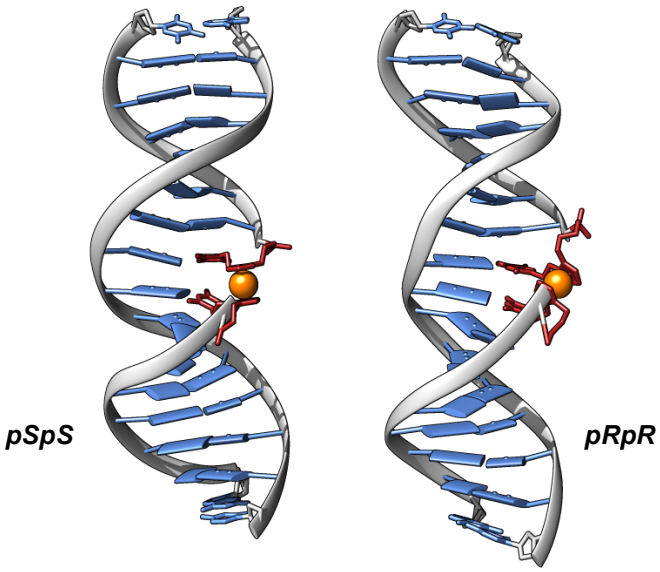


*Figure 3.30 - MD simulations (6 ns) for **Fcn3c2TT*pYpY*** when incorporated into **S1** and duplexed with **S2AA**.*

3.2.3.5 Decreasing the Number of Carbons to the Phosphate.

Next it was decided to change the number of carbons in the phosphate linker. However, due to constraints in the synthesis of the monomers it had to be designed without the methyl at the C1 position. For this system only the ***pRpR*** isomer was synthesised, along with the abasic control (which has no chirality), as it most likely mimicked the helical twist of the more successful **Fcn3c2TT*RRpSpS*** motif (Figure 3.31).

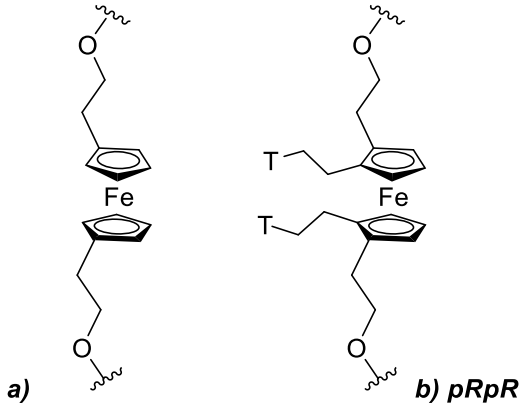


*Figure 3.31 - The structure of FcNA monomer **Fcn2c2TT** showing the b) **pRpR** stereoisomer synthesised and the abasic control a) **Fcn2c2HH**.*

The T_m values (Table 3.6) show that reduction in the carbon linkage results in a reduction in stability (-1.0 °C compared with **Fcn3c2TT**$_{RRpSpS}$) while keeping the discrimination between opposing nucleobases AA and GG (42 °C vs 40.5 °C respectively). These results can be explained by the removal of the favourable hydrophobic interactions arising from the methyl group and a reduction in linkage flexibility imposing more strain on the duplex overall.

*Table 3.6 – T_m values for **Fcn2c2TT**$_{pRpR}$ and abasic monomer **Fcn2c2HH** in **S1** when duplexed with **S2YY**. 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

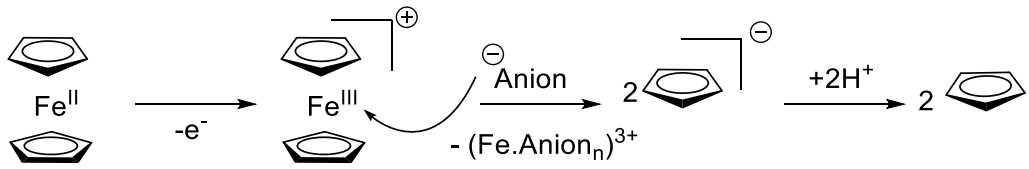
	Thermal Melting Temperature (T_m) / °C		
Strand	**S2AA**	**S2YY= S2GG**	**S2TT**
S1Fcn2c2HH	36.5 (±0.2)	37.0 (±0.0)	35.5 (±0.3)
S1Fcn2c2TT$_{pRpR}$	42.0 (±0.0)	40.5 (±0.2)	41.0 (±0.0)
S1TT	56.0 (±0.2)	44.0 (±0.0)	41.5 (±0.2)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

3.2.4 The Stability of FcNA

All of the T_m results described above were carried out in a phosphate buffered solution (pH 7.0) containing 100 mM $NaClO_4$. However normally NaCl is used as the salt in these experiments as it more closely mimics natural conditions. The first T_m measurements were in fact performed using NaCl and similar, albeit slightly higher T_m values were observed for the

ferrocene modified strands. However, it was not until a closer inspection of the stability of these modified strands was carried out that stability problem was identified. Originally, the purified strands were stored in a freezer and thawed prior to use and sometimes these samples were left overnight at room temperature. Analytical HPLC runs of the stocks showed the emergence of additional peaks in all the samples that eluted earlier than the desired products (Figure 3.32). Degradation of ferrocene derivatives in aqueous media is not unusual[97-99] and has been reported for ferrocene labeled DNA by Plaxco and co-workers.[100] Degradation of ferrocene under aqueous conditions has been postulated as occurring *via* destruction of the oxidized form of the ferrocene structure, ferrocenium (Fc^+), upon attack of nucleophilic anions (Scheme 3.1).[101]



Scheme 3.1 - Nucleophilic anion attack and degradation of ferrocenium species.

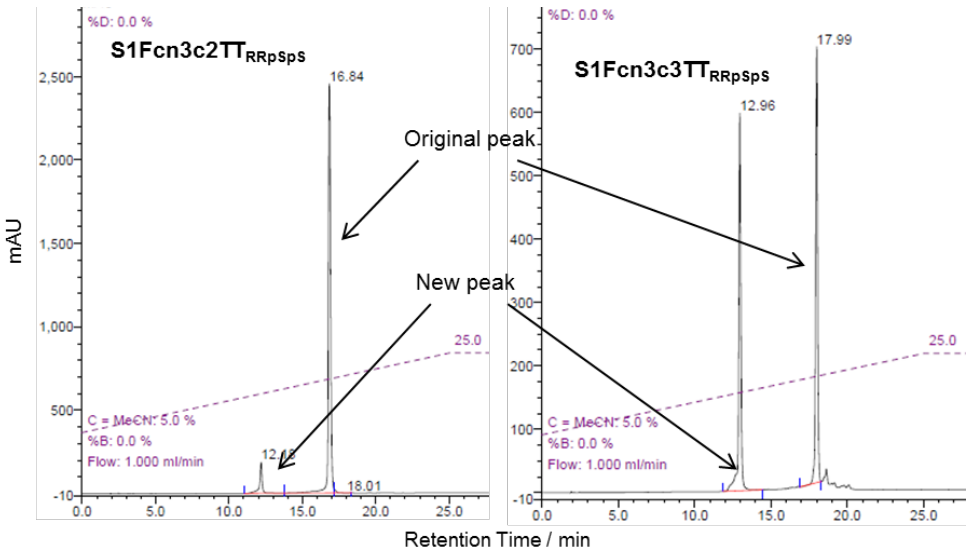


*Figure 3.32 - Analytical HPLC traces of two strands of FcNA modified **S1** after overnight exposure to room temperature.*

Initially the additional peak in the HPLC was attributed to degradation products. However, upon isolation of the peak and further analysis of these peaks by mass spectrometry the picture became less clear. For all four variants of FcNA (and each stereisomer) the additional peak produced a molecular ion peak in the mass spectrum which showed the same mass as the intact and pure synthesised oligonucleotides.

It was postulated that the ferrocene unit itself was oxidising in solution to produce a ferrocenium (Fc^+) derivative of each oligonucleotide that was stable enough to be isolated. This was confirmed by investigating the UV-vis spectrum of concentrated (100 μM) aliquots of both the original fraction and the newly emerged peak (Figure 3.32) for **S1Fcn3c2TT*RRpSpS*.** Ferrocene shows a band centred at 450 nm ascribed to its d-d transition; this is shown for the original peak (Figure 3.33a). The proposed Fc^+ species gives rise to a band centred at 680nm. By comparing the spectra of these two peaks with that of ferrocene and the ferrocenium tetrafluoroborate (BF_4) salt (Figure 3.33b), the similarities are clear, with ferrocene displaying a d-d band centred at 440 nm and the salt a band at 620 nm. Furthermore, the DNA solutions appeared different to the naked eye, the original stock (100 100 μM) having a slight tinge of yellow/orange and the isolated 'degradation' product appearing blue.

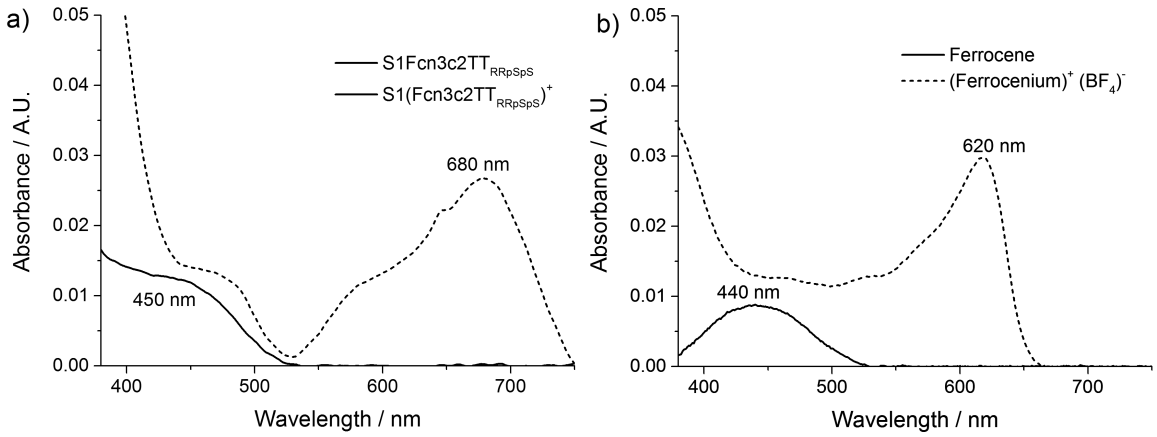


Figure 3.33 - UV-vis spectra comparing ferrocene and ferrocenium derivatives of a) ***S1*** *and b) underivatised molecules. Concentration of each compound is 100 μM. DNA and ferrocenium tetrafluoroborate measured in 10 mM sodium phosphate buffer pH 7.0, 10 mM $NaClO_4$, ferrocene measured in acetonitrile.*

It then became of interest to investigate whether the oxidised forms of the oligonucleotides caused a change in overall duplex stability. It was at this point decided that the sodium salt should be changed from chloride to perchlorate for all further experiments. Studies have previously shown that nucleophilic anions such as Cl^- can cause degradation of Fc^+ derivatives, with the rate of degradation a function of the nucleophilicity of the anion (Cl^- > Br^- > NO_3^- > SO_4^{2-} > ClO_4^-).[97,99] At this stage, all oligonucleotide stocks were stored in degassed ultra-pure Milli-Q water (at -20 °C) in order to reduce the possible oxidation of stock solutions and thus aliquoted into appropriate volumes to minimize the exposure to elevated temperatures; these procedures and conditions were those recommended in a number of papers[100,101] and were successful in reducing the generation of the oxidised strand forms in the stock solutions.

Interestingly, when comparing the duplex stabilities of oxidized and unchanged oligonucleotides with **S2AA**, a difference in the T_m was observed (Table 3.7), within all cases, the oxidized form giving an increase in the overall stability. Strand **S1Fcn3c3TT*$_{SSpRpR}$*** shows the largest stability enhancement (+2.5 °C). This increase in stability is not unexpected; by

introducing a positive charge to the Fc unit a reduction in the overall negative charge on the backbone of each strand would be expected. Furthermore, the positive charge would make the organometallic unit less hydrophobic in character, which would lower the competition with the nucleobases for the hydrophobic DNA core. These findings show great promise for future control of duplex stability through controlled chemical or electrochemical oxidation of oligonucleotides.

*Table 3.7 - T_m values comparing the ground state and oxidised form of FcNA monomers incorporated into **S1** when duplexed with their full complement **S2AA**. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

		T_m / °C **S2AA**	
S1XX	Fc	Fc^+	($\Delta T_m = Fc^+ - Fc$)
Fcn3c2TT$_{RRpSpS}$	43.0 (±0.1)	44.5 (±0.4)	+1.5
Fcn3c2TT$_{SSpRpR}$	41.0 (±0.3)	41.5 (±0.3)	+0.5
Fcn3c3TT$_{SSpRpR}$	43.0 (±0.1)	45.5 (±0.4)	+2.5
Fcn3c3TT$_{RRpSpS}$	41.5 (±0.2)	42.5 (±0.6)	+1.0
Fcn3c2TT$_{pSpS}$	40.5 (±0.2)	41.0 (±0.3)	+0.5
Fcn3c2TT$_{pRpR}$	41.5 (±0.0)	42.0 (±0.0)	+0.5
Fcn2c2TT$_{pRpR}$	42.0 (±0.0)	42.5 (±0.0)	+0.5

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

3.2.5 FcNA Pyrimidines

Changes to the phosphate and nucleobase linkers attached to the FcNA (TT) monomer units results in subtle differences in duplex melting temperatures. It was evident that a methyl group at the α position is important and arguments for using either C2 and C3 linkers to the nucleobase can be made. However due to the apparent greater chemical stability of the C2 monomer (**Fcn3c2TT$_{XXpYpY}$**), this linker length was chosen for further studies involving other nucleobases appended to the Cp ring as described below. *Please note that a full and comprehensive list of all of the discussed T_m values, as well as other controls are included in appendices 4-7.

Herein the monomeric design **Fcn3c2TT$_{XXpYpY}$** is described more simply as **FcBB$_{XX}$**; where **BB** represents the appended nucleobase and ***XX*** the chirality about the methyl

position. As described earlier, in this system the planar chirality is the opposite of the central chirality i.e. ***RR*** results in ***pSpS***.

3.2.5.1 Comparing thymine and cytosine

Given that the incorporation of thymine bearing ferrocene monomers results in a significant loss in overall duplex stability over unmodified DNA (>13 °C), and the difference in T_m between complementary (**S2AA**) and double mismatch (**S2GG**) strands was clearly not as large (ΔT_m = -1.5 °C for **S1FcTT*RR*** and ΔT_m = -11.5 °C for **S1TT**), it was anticipated that appending other nucleobases to the monomer unit might result in greater stability effects. Cytosine for example, forms three complementary hydrogen bonds with guanine, and was therefore expected to result in higher T_m values.

Due to the synthetic challenges associated with making these monomers, the ***RR*** stereoisomer of a cytosine monomer was first synthesised; this isomer gave the highest stabilities for thymine systems. This monomer was accordingly incorporated into **S1** *via* automated oligonucleotide synthesis, as described previously.

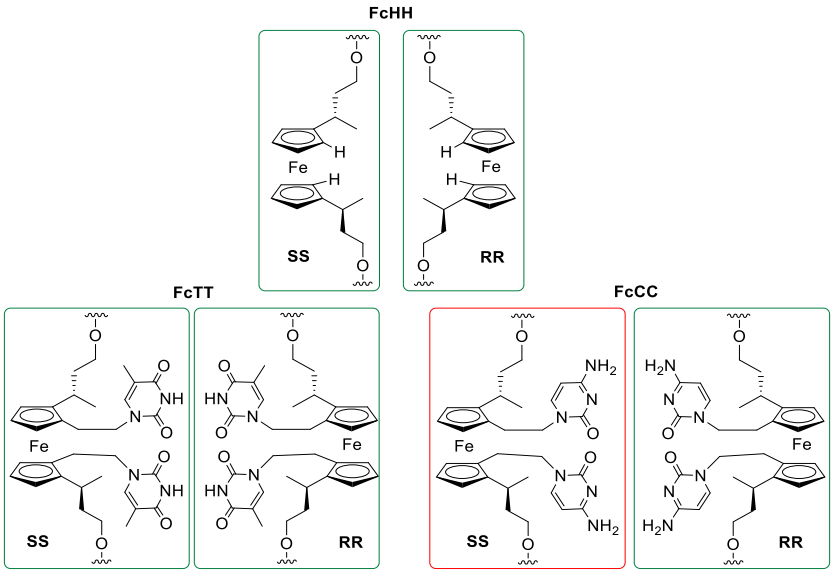


*Figure 3.34 – The FcNA (**Fcn3c2BB**) monomers incorporated into **S1** bearing pyrimidines thymine and cytosine (green). The **SS** isomer of **FcCC** is still to be synthesised and incorporated.*

Thermal melting analysis showed promising and significant trends for this system (Table 3.8). The cytosine strand **S1FcCC$_{RR}$** when duplexed with its fully complementary strand gives a T_m value 5 °C higher than for the equivalent thymine system bound to **S2AA**. This increase is infact similar to that found for the unmodified analogues (60 °C and 56 °C respectively. Furthermore, the difference in T_m between the complementary (**S2GG**) and double mismatch (**S2AA**) strands is much greater, with values of 48 °C and 39.5 °C respectively. The extent of this discrimination can be visualised through the melting profiles shown in Figure 3.35. Furthermore, the target offering only one possible H-bonding base pair (**S2AG**) has an intermediate melting temperature of 45.5 °C. Target **S2TT** produces the lowest T_m, which can be explained by a loss of base stacking contributions as well as a H-bonded base pair, as would be expected. The equivalent experiments with RNA and PNA targets (Table 3.9 and Table 3.10 respectively) show the same trends, but as expected, the discrimination is even greater owing to the enhanced binding strength of DNA•RNA and DNA•PNA duplexes over DNA•DNA sytems.

*Table 3.8 - Thermal Melting Temperatures (T_m) of FcNA (**Fcn3c2BB$_{RRpSpS}$**) incorporations into **S1XX** with targets **S2YY**. (BB = HH, TT or CC). 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C* S2YY=				
	S2AA	**S2AG**	**S2GG**	**S2TT**	**S2AbAb**
FcHH$_{RR}$	37.5 (±0.2)	37.5 (±0.3)	37.5 (±0.3)	36.0 (±0.0)	35.5 (±0.3)
FcTT$_{RR}$	43.0 (±0.1)	42.0 (±0.0)	41.0 (±0.0)	40.5 (±0.2)	38.5 (±0.2)
FcCC$_{RR}$	39.5 (±0.3)	45.5 (±0.3)	48.0 (±0.3)	37.5 (±0.2	37.0 (±0.0)
AbAb	37.0 (±0.0)	37.0 (±0.0)	37.0 (±0.0)	34.0 (±0.0)	--
TT	56.0 (±0.2)	33.0 (±0.3)	44.0 (±0.0)	41.5 (±0.2)	33.0 (±0.3)
CC	38.0 (±0.0)	50.5 (±0.3)	60.0 (±0.0)	38.0 (±0.0)	32.5 (±0.5)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

*Table 3.9 - Thermal Melting Temperatures (T_m) of FcNA (**Fcn3c2BB$_{RRpSpS}$**) incorporations into **S1XX** with RNA targets **S2YYR**. (BB = HH, TT or CC) (RNA targets). 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C*		
		S2YY=	
S1 XX =	**S2AAR**	**S2GGR**	**S2UUR**
FcHH$_{RR}$	39.0 (±0.0)	36.0 (±0.0)	35.0 (±0.0)
FcTT$_{RR}$	47.5 (±0.3)	43.0 (±0.0)	40.0 (±0.0)
FcCC$_{RR}$	42.5 (±0.5)	49.5 (±0.3)	39.5 (±0.3)
AbAb	40.0 (±0.0)	38.0 (±0.0)	33.0 (±0.6)
TT	59.0 (±0.0)	52.5 (±0.3)	44.5 (±0.1)
CC	45.0 (±0.0)	69.5 (±0.3)	39.0 (±0.0)

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

*Table 3.10 - Thermal Melting Temperatures (T_m) of FcNA (**Fcn3c2BB$_{RRpSpS}$**) incorporations into **S1XX** with targets **S2YYP**. (BB = HH, TT or CC). 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$. (A.S data)*

	Thermal Melting Temperature (T_m) / °C*			
		S2YY=		
S1 XX =	**S2AAP**	**S2AGP**	**S2GGP**	**S2TTP**
FcHH$_{RR}$	60.0	59.0	59.0	55.0
FcTT$_{RR}$	67.0	64.5	62.0	59.0
FcCC$_{RR}$	61.5	64.0	71.0	57.0
AbAb	58.0	57.0	55.0	52.0
TT	81.0	72.0	72.5	65.5
CC	58.5	69.5	86.0	54.0

*Highlighted values represent 'fully complementary' sequences. Error in instrument data = ±0.5 °C.

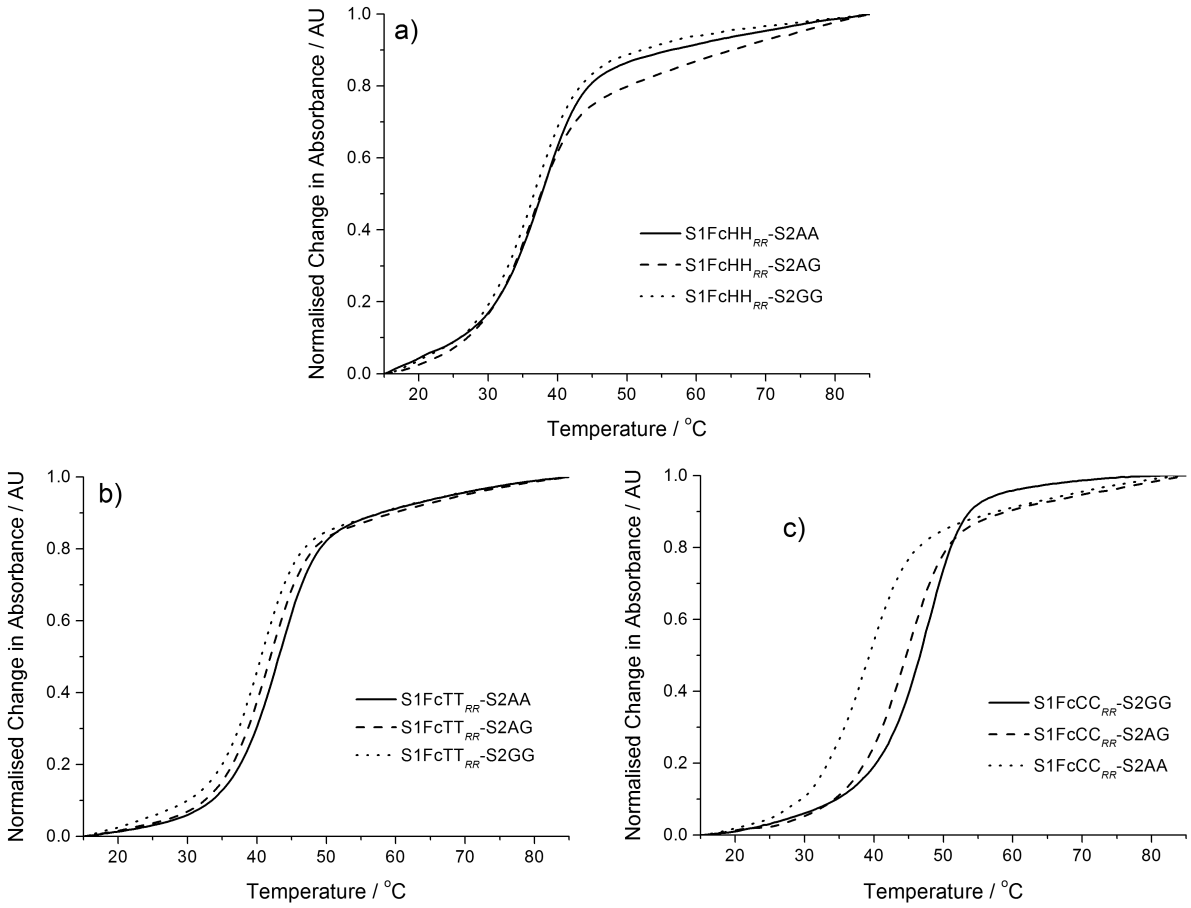


*Figure 3.35 – Normalised (@260 nm) thermal melting profiles of strands **S1** with FcNA monomers incorporated into S1 a) **FcHH$_{RR}$** b) **FcTT$_{RR}$** c) **FcCC$_{RR}$**, when duplexed with **S2YY** (where **YY** = AA, AG or GG). 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

Determination of the thermodynamic parameters for these complexes via van't Hoff analysis also show the same trends (Table 3.11) with more negative *ΔG* values at 293 K found for those duplexes with the higher T_m values. However some caution has to be taken when considering these values, these calculations make some assumptions about the bimolecular nature of the interactions as well as the the enthalpy of the system (discussed in Chapter 2). The *ΔH* and *ΔS* values determined from these plots are detailed in appendix 7.

Table 3.11 - Gibbs free energy values at 293K calculated from van't Hoff analysis of the thermal melting curves. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.

	$\Delta G_{20°C}$ / kJ mol^{-1}		
S1 XX =	**S2AA**	**S2AG**	**S2GG**
$FcHH_{RR}$*	-49.0 (±2.0)*	-48.5 (±1.7)*	-49.2 (±1.5)*
$FcTT_{RR}$	-60.2 (±0.9)	-58.5 (±0.6)	-58.1 (±0.8)
$FcCC_{RR}$	-61.1 (±2.6)	-69.7 (±0.4)	-72.5 (±1.0)
TT	-82.9 (±1.9)	-68.5 (±1.9)	-62.5 (±1.1)
CC	-53.3 (±0.8)	-76.8 (±0.6)	-88.6 (±2.4)

*Inaccuracies in lower baseline determination due to transition starting at low temperatures. (.) show the standard error of the mean of a minimum of three experimental repeats. Parameters calculated using van't Hoff plots according to the procedure detailed in Chapter 2.2.2.

3.2.5.2 Circular Dichroism Spectroscopy

CD spectroscopy was used to study the two pyrimidine systems (thymine and cytosine), and the abasic control, while also comparing FcNA modified sequences with unmodified (Figure 3.36). The **$S1FcHH_{RR}$** system was found to have a very similar CD spectrum to that of its unmodified DNA equivalent (**S1AbAb**) when bound to targets **S2AA** and **S2GG** (Figure 3.36 a,b). The fact that the T_m values for these systems are very similar and the similarity of the CD spectra suggest that the incorporation of ferrocene has a negligible effect on the bulk B-DNA structure when no bases are appended to it.

For the thymine derivative **$S1FcTT_{RR}$** the differences are more pronounced. For example, the B-DNA bands at 275 and 248 nm are weaker in intensity, suggesting an untwisting of the helical stack (Figure 3.36d). Interestingly when duplexed with **S2GG** (double mismatch) the B-DNA bands are slightly more intense, which is not the case for its unmodified equivalent (Figure 3.36c). This suggests that in order for the rigid ferrocene unit to participate in base-pairing with the two adenines opposite, the duplex must compensate by distorting or unwinding to some extent. This would support the dramatic difference in T_m between modified (**$S1FcTT_{RR}$•S2AA** = 43 °C) and unmodified duplexes (**S1TT•S2AA** = 56 °C). Another difference is that the positive band at 218 nm appears more intense for the ferrocene systems relative to the main positive band at *ca.* 275 nm. Ferrocene does absorb strongly at this wavelength and the increase in intensity for the fully complementary target

suggests that the ferrocene unit is in a more chiral environment in this duplex. Molecular dynamic simulations with double mismatching targets are planned for future studies in order to postulate a reasoning for these observed differences.

The cytosine derivative **S1FcCC$_{RR}$** displays the largest spectral differences between modified and unmodified strands, being the most pronounced for the complementary and double mismatch targets (Figure 3.36e,f). The change from target **S2AA** to **S2GG** results in a broadening and blue shift of the two characteristic B-DNA bands.

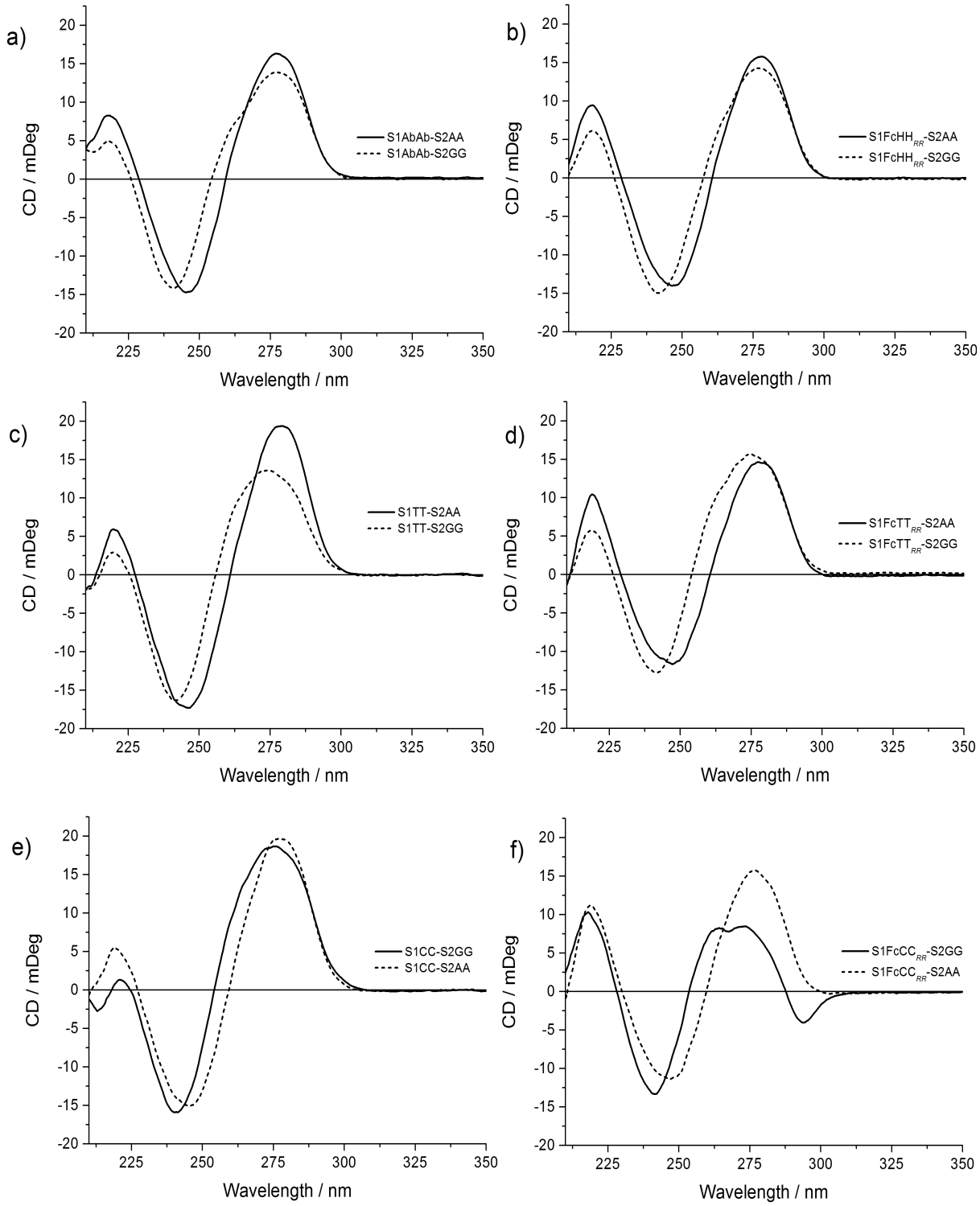


Figure 3.36 - CD spectra for a) abasic controls ***S1AbAb$_{RR}$•S2YY*** *and b)* ***S1FcHH$_{RR}$•S2YY*** *(where* ***YY*** *= AA (solid) and GG (dashed).* ***S1XX*** *sequences and unmodified controls when bound to complementary (solid line) and mismatching (dashed line) strands for where* ***XX*** *= c) TT d)* ***FcTT$_{RR}$*** *e) CC f)* ***FcCC$_{RR}$****. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

However, for **S1FcCC$_{RR}$** an additional change is observed, consisting of a large decrease in the intensity of the positive band, and addition of a shoulder, and most notably, the emergence of a negative band at 290 nm. The large change or disruption to the B-DNA structure for this particular **FcCC$_{RR}$** system is likely to be a consequence of the increased rigidity afforded by the formation of two relatively strong, compared to the weaker **FcTT$_{RR}$,** C-G base pairs. As a result, the duplex might be expected to loosen and perhaps even flip a base out of the helical twist, which could explain the negative band observed. Therefore differences in both the T_m data and CD spectra for this **FcCC$_{RR}$** system definitely provide good evidence for the formation of unnatural-natural cross duplex interactions. Modelling studies that could support these findings are currently underway.

3.2.6 FcNA Purines

With there now being good evidence for Watson-Crick H-bonding between FcNA units and DNA bases, it was decided to study all four nucleobases in the same strand sequence for a complete set. Accordingly stereoisomers of the adenine derivative were synthesised and incorporated into the S1 strand. The guanine derivatives were found to be significantly more challenging to synthesise owing to difficulties with solubility and protection strategies, and this work is ongoing in the group.

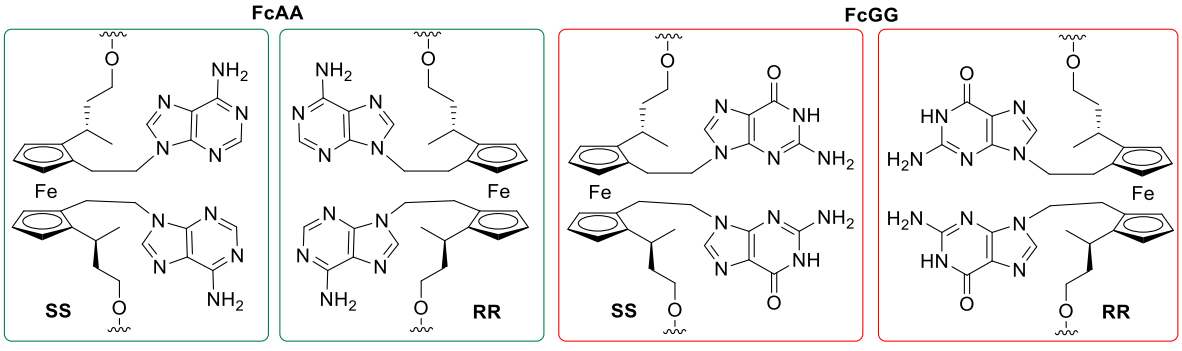


*Figure 3.37 – The synthesised monomers of **FcAA** incorporated into **S1** and the guanine equivalents that are yet to synthesised.*

For this system, the ***SS*** stereoisomer produced the highest duplex stability with its complementary target **S2TT** (47 °C, Table 3.12). Compared with the unmodified control (**S1AA•S2TT**) this gave a 10.5 °C decrease in overall stability, but a significant increase compared to the analagous thymine system (43 °C, Table 3.8). The change in preference for ***SS*** over ***RR*** is difficult to explain but is probably associated with enhanced stacking interactions. In fact this system could participate in stabilising 'zipper-like' interactions[102,103] as a result of an increase in the overlap of flanking nucleobases (i.e. CTTC•GAAG vs CAAC•GTTG). When bound to **S2CC** (CAAC•GCCG) the T_m is considerably lower, showing a large discrimination (8.5 °C) between complementary and double mismatch targets. These differences are highlighted in Figure 3.38, with pyrimidine bases in the target giving the highest (TT, H-bonding) and lowest (CC, weakest π-stacking) values.

The same discrimination behaviour is not observed for the ***RR*** stereoisomer, giving a lower duplex stability for the complementary target (**S2AA**) and fairly uniform values for all targets, although the double abasic (**S2AbAb**) is the lowest. This suggests that the cavity is too small or too conformationally strained to easily accommodate the FcNA unit with this stereochemistry. As a result the duplex could be unwinding or buckling the flanking nucleobases, or even unable to form hydrogen bonds at all. MD simulations of both stereoisomers are on-going, and are expected to provide an insight into the cause of these observations.

*Table 3.12 – T_m values for FcNA purine modifications for the duplex **S1XX•S2YY**. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C				
S1 XX =	**S2TT**	**S2CC**	**S2GG**	**S2AA**	**S2AbAb**
$FcHH_{RR}$	36.5 (±0.2)	36.0 (±0.3)	37.5 (±0.3)	38.0 (±0.2)	35.5 (±0.5)
$FcHH_{SS}$	36.0 (±0.2)	35.5 (±0.3)	41.0 (±0.3)	37.5 (±0.2)	36.0 (±0.0)
$FcAA_{RR}$	44.5 (±0.3)	43.0 (±0.3)	42.5 (±0.9)	42.0 (±0.3)	39.5 (±0.3)
$FcAA_{SS}$	47.0 (±0.0)	38.5 (±0.1)	41.5 (±0.3)	41.0 (±0.0)	39.0 (±0.0)
$FcGG_{RR}$	TBD	TBD	TBD	TBD	TBD
$FcGG_{SS}$	TBD	TBD	TBD	TBD	TBD
AA	57.5 (±0.2)	38.5 (±0.3)	46.0 (±0.2)	46.5 (±0.3)	37.5 (±0.3)
GG	TBD	TBD	TBD	TBD	TBD

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

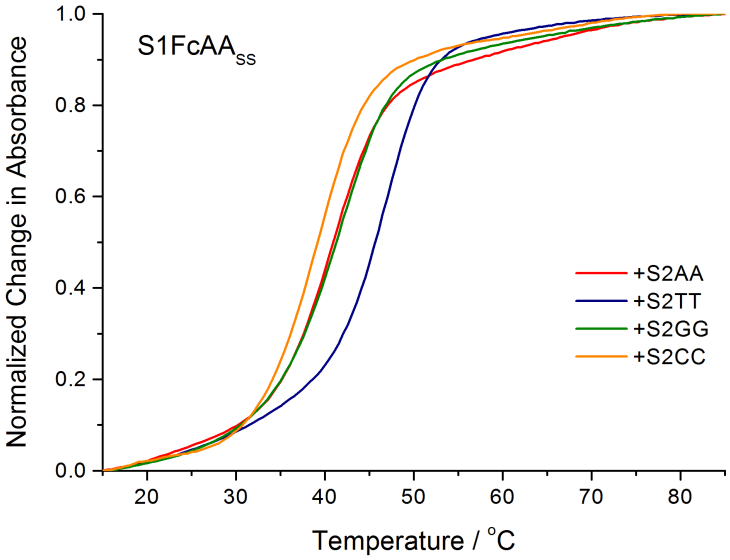


*Figure 3.38 - Normalised $_{(@260\ nm)}$ T_m curves for **S1FcAA$_{ss}$** when duplexed with **S2YY** (**YY** = AA,TT,GG,CC) showing the base discriminating behaviour. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

As done previously with TT and CC systems, interactions with other nucleic acid targets RNA (Table 3.13) and PNA (Table 3.14) were next investigated. Once again the trends were replicated, with the ***SS*** stereoisomer giving the highest stability with fully complementary RNA and PNA strands.

*Table 3.13 - T_m values for FcNA purine modifications for the duplex **S1XX•S2YYR**.(RNA targets). 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C			
S1XX =	**S2UUR**	**S2CCR**	**S2GGR**	**S2AAR**
$FcHH_{RR}$	35.0 (±0.0)	34.5 (±0.3)	36.0 (±0.0)	39.0 (±0.0)
$FcHH_{SS}$	35.0 (±0.0)	34.5 (±0.2)	36.0 (±0.6)	39.5 (±0.3)
$FcAA_{RR}$	44.0 (±0.0)	40.5 (±0.4)	TBD	43.5 (±0.6)
$FcAA_{SS}$	50.5 (±0.3)	43.0 (±0.3)	TBD	43.0 (±0.0)
$FcGG_{RR}$	TBD	TBD	TBD	TBD
$FcGG_{SS}$	TBD	TBD	TBD	TBD
AA	55.0 (±0.0)	42.5 (±0.5)	46.0 (±0.0)	42.5 (±0.5)
GG	TBD	TBD	TBD	TBD

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

*Table 3.14 – T_m values for FcNA purine modifications for the duplex **S1XX•S2YYP** (PNA targets). 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$. (A.S data)*

	Thermal Melting Temperature (T_m) / °C		
S1XX =	**S2TTP**	**S2GGP**	**S2AAP**
$FcHH_{RR}$	55.0	59.0	60.0
$FcHH_{SS}$	TBD	62.0	60.0
$FcAA_{RR}$	61.0	61.5	61.5
$FcAA_{SS}$	69.5	59.0	59.0
$FcGG_{RR}$	TBD	TBD	TBD
$FcGG_{SS}$	TBD	TBD	TBD
AA	74.0	64.5	64.5
GG	TBD	TBD	TBD

*Highlighted values represent 'fully complementary' sequences. Error in instrument data = ±0.5 °C.

CD spectra of the most stable motif (**S1$FcAA_{SS}$**) with the fully complementary (**S2TT**) and double mismatch (**S2CC**) strands shows only subtle differences in the B-DNA bands (Figure 3.39a) with no evidence for a large change in base stacking interactions. The major differences comes from the band at 218 nm, which is made more intense by the ferrocene π-π* transition and is even more pronounced fror the complementary strand. This could be a result of the ferrocene unit being incorporated into the twist of the DNA helix more tightly as a result of the H-bonded base pairing.

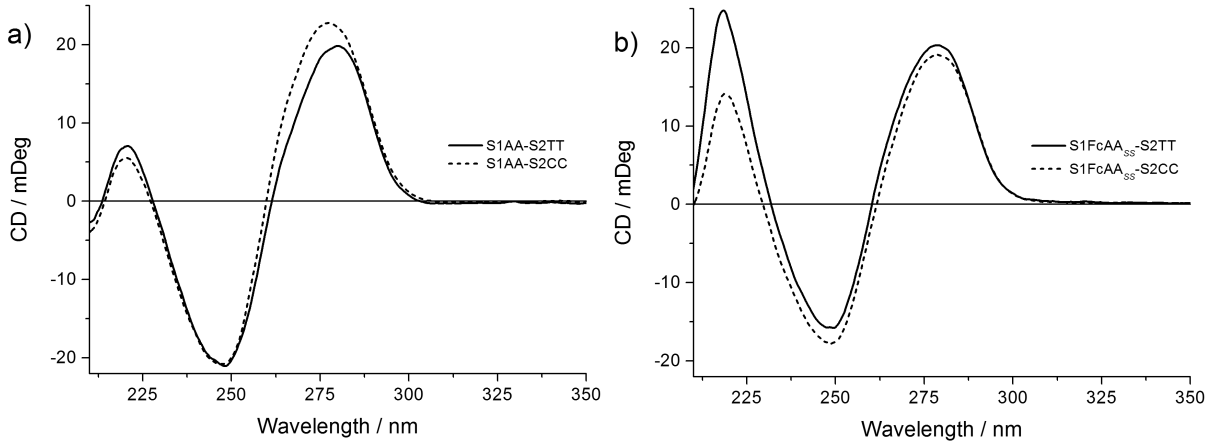


*Figure 3.39 - CD spectra for a) **S1AA•S2YY** and b) **S1FcAA$_{SS}$•S2YY** (where **YY** = TT (solid) and CC (dashed). 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

3.2.7 Electrochemistry

The incorporation of FcNA units into oligonucleotides gives the possibility for electrochemical applications in sensing and nanoelectronics due to the reversible redox behaviour of ferrocene. FcNA monomers have previously been shown to be redox-active[91] as well oligomeric octamers systems.[89,104] It was therefore important to assess the electrochemical behaviour of these units when incorporated into DNA.

3.2.7.1 Electrochemical Behaviour

The four most promising FcNA-DNA conjugate strands (**S1FcHH$_{RR}$**, **S1FcTT$_{RR}$**, **S1FcCC$_{RR}$** and **S1FcAA$_{SS}$**), in terms of the highest duplex stability and base pairing behaviour were shown to be redox-active in solution (50 μM). This concentration was used to allow for the discrimination of electrochemical outputs from background capacitive currents. Cyclic voltammetry (discussed in Chapter 2) was used to probe their behaviour, where the labelled oligonucleotides were interrogated with a scanning/ramped potential between -0.1 to +0.35 V at different scan rates. The cyclic voltammograms of each strand are shown in Figure 3.40 alongside their respective Randles-Ševčik plots of the measured currents. The scan rate

dependence for each system shows a linear relationship with the square root of the scan rate, which is indicative of diffusing redox-active species. Diffusion coefficients *D* (shown in Table 3.15) were calculated by taking the slope of these linear plots according to the Randles-Ševčik equation (Chapter 2 – Equation 2.8).[105,106]

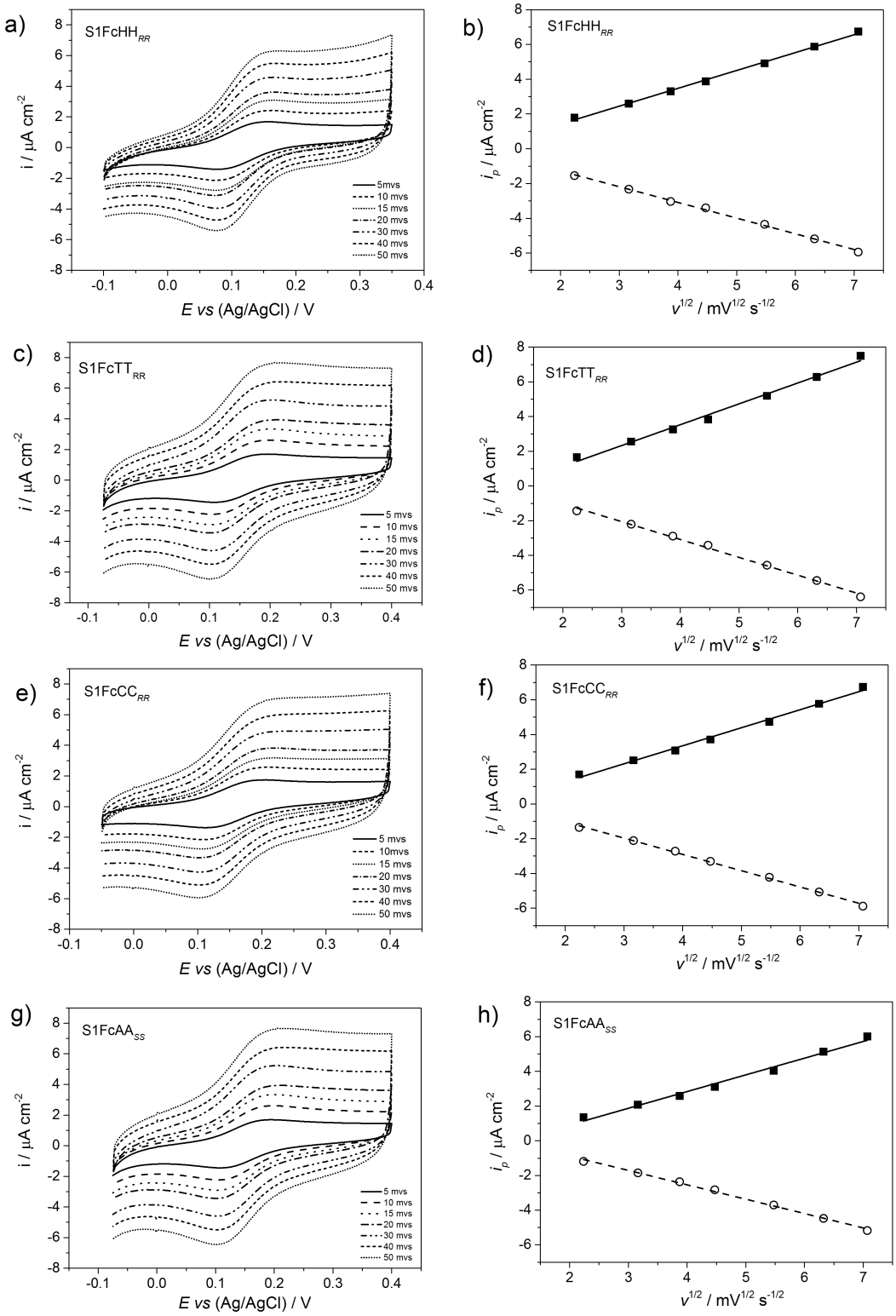


Figure 3.40 - Cyclic voltammograms at increasing scan rates (left) and the corresponding Randle-Sevcik plots (right) of the observed current as a function of the square root of the scan rate for (a,b) ***S1FcHH$_{RR}$****, (c,d)****S1FcTT$_{RR}$****, (e,f)****S1FcCC$_{RR}$*** *and (g,h)****S1FcAA$_{SS}$****.■ – anodic current,○ – cathodic current. 50 µM DNA in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$.*

The observed behaviour (data in Table 3.15) for each strand indicates reversible Nernstian behaviour ($\Delta E_p \approx 59$ mV), with strands **S1FcHH$_{RR}$** and **S1FcTT$_{RR}$** ($\Delta E_p = 62$ mV) seemingly more reversible than the cytosine ($\Delta E_p = 75$ mV) and adenine ($\Delta E_p = 77$ mV) counterparts. The cathodic to anodic peak ratio (i_p^c/i_p^a) is less than one for all four systems, however this is likely a result in the difference in diffusivity of the duplexes upon oxidiation, brought about by a change in overall charge.

Table 3.15 - Electrochemical behaviour of FcNA labelled oligonucleotide strand S1.(v = 15 mv s^{-1} for CV and 1 mV step, 25 mV amplitude and a 8 Hz frequency for SWV). 50 µM DNA in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$. Potentials are reported versus Ag/AgCl.Confidence limit is ±5 mV for CV and ±2 mV for SW.

		CV			SWV
Strand	$\Delta E_p = E_p^a - E_p^c$ / mV	$E_{1/2}$ / mV	i_p^c/i_p^a	D / cm^2 s^{-1}	$E_{1/2}$ / mV
S1FcHH$_{RR}$	+62	112	0.89	2.42 x10^{-9}	113
S1FcTT$_{RR}$	+62	148	0.98	2.82 x10^{-9}	147
S1FcCC$_{RR}$	+75	155	0.86	2.44 x10^{-9}	155
S1FcAA$_{SS}$	+77	173	0.89	2.26 x10^{-9}	175
S1FcGG$_{SS}$	TBD	TBD	TBD	TBD	TBD

The $E_{1/2}$ value for each strand gives information regarding the ease of their oxidation and the data indicates quite a wide range of values (Table 3.15). Overlays of the CVs and SWV's for the four systems are shown in Figure 3.41.

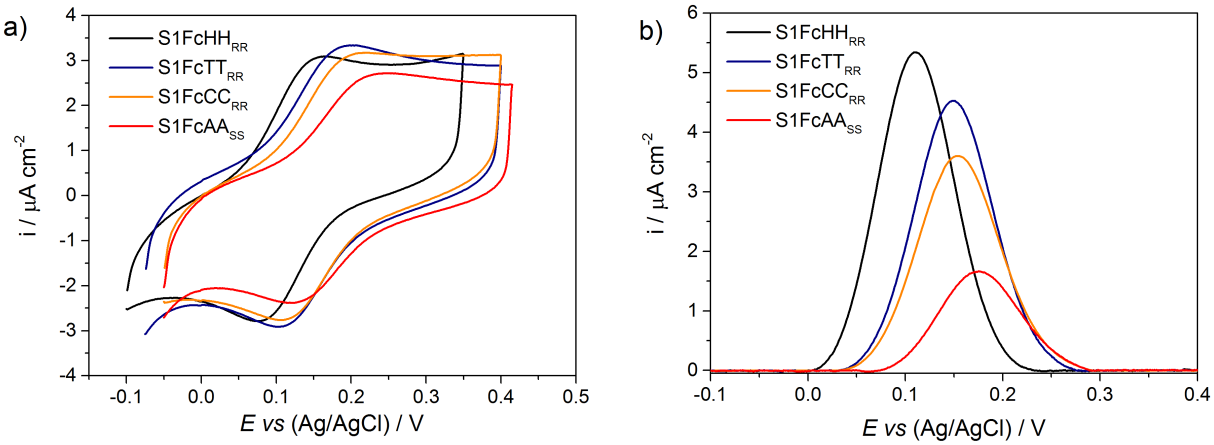


Figure 3.41 - a) CV (v = 15 mv s^{-1}) and b) SWV overlays of the four FcNA modified strands. 50 µM DNA in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$.

The reasoning for these differences is attributable to differences in the electron-withdrawing effects of the appended nucleobase. As expected **S1FcHH$_{RR}$** has the least positive $E_{1/2}$ potential (112 mV) as it remains unaffected by appended electron deficient nucleobases. The difference in $E_{1/2}$ for thymine (148 mV) and cytosine (155 mV) follows the same trend as the relative electron affinity energies of the nucleobases calculated by Dutta *et al.* (G<T<C<A).[107] Predictably **S1FcAA$_{SS}$** requires the most positive potential (173 mV) to oxidise ferrocene attributable to the more electron deficient purine rings that are pulling electron density away from ferrocene. Accordingly, it might be expected that guanine FcNA systems would have the lowest oxidation potential. In line with these results, Houlton and co-workers have previously shown the oxidation potential of ferrocenyl-nucleobases to be dependent on the electron withdrawing abilities of each nucleobase (Figure 3.42).[81] 9-ferrocenyl-methyladenine (**3.20**) required the highest oxidation potential to remove and electron compared with a cytosine (**3.19**) derivative, whilst substitution of the C6 amino functionality of adenine for Cl (**3.21**) resulted in the most positive potential.[81]

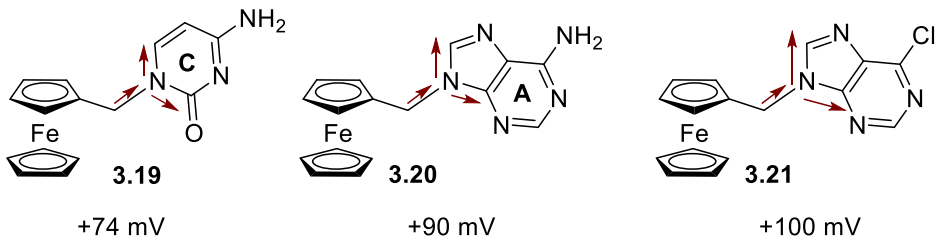


Figure 3.42 - Electron withdrawing effects of 9-ferrocenyl-methylcytosine (3.19), 9-ferrocenyl-methyladenine (3.20) and adenine derivative 2-amino-6-chloro-9-ferrocenyl-methylpurine (3.21) highlighting the oxidation potentials (vs. Fc/Fc+) in acetonitrile, as reported by Houlton et al.[81] Scan rate of 500 mv s^{-1}.

Similarly previous work in the Tucker group has shown bis-substituted ferrocene derivatives to have more positive oxidation potentials when adenine (**3.23**) is appended and compared with thymine (**3.22**, Figure 3.43). Furthermore, the addition of electron deficient benzoyl groups to the heterocycles resulted in a overall further shift to more positive

oxidation potentials; whilst adenine (**3.26**) again produced the highest oxidation potentials when compared with cytosine (**3.24**) and uracil (**3.25**) derivatives. Uracil likely falls between cytosine and adenine due to the loss of the +I effect of the methyl group present in thymine.

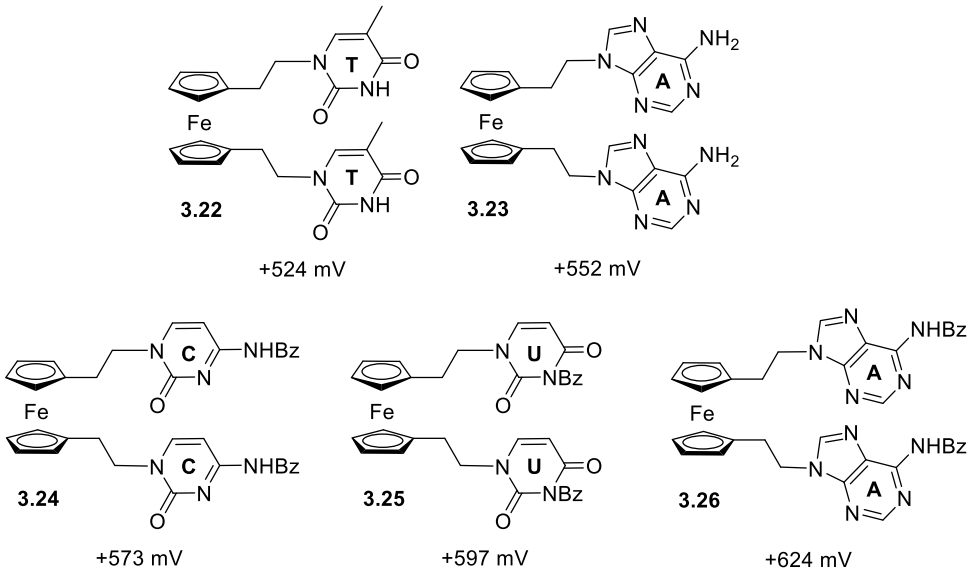


*Figure 3.43 - Oxidation potentials of bis-substituted ferrocenes with nucleobases without (**3.22, 3.23**)[88] and with (**3.24-3.26**)[91] benzoyl (Bz) protecting groups (in DCM with 0.1 mM tetrabutylammoniumhexafluorophosphate). Potentials are reported vs. dmfc which was used as an internal reference.*

3.2.7.2 Addition of Targets

Initial screening of these nucleobase modified systems upon the addition of DNA targets **S2YY** to the strands showed some discriminative behaviour between complementary and double mismatch targets for the thymine derivatives. The observed current decreased significantly in both the CV's and SWV's and a small shift of the $E_{1/2}$ (SWV) to a less positive potential was observed in all cases (Table 3.16). Large current reductions would be expected regardless of the base opposite in the sequence due to the change in bulk properties (size, charge and diffusivity) of the labelled strands. The shift to less positive potential is also expected as the extra negative charge from the target can stabilise the ferrocenium ion.[62] The ability to discriminate could be the result of subtle differences in the properties or as a result of probing the direct mismatch/match environment in which the FcNA label sits.

*Table 3.16 - The change in electrochemical behaviour upon addition of complementary targets to FcNA modified **S1**. 50 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$. Potentials are reported versus Ag/AgCl.Confidence limit is ±2 mV for SW.*

Strand	Change in $E_{1/2}$ / mV	% Current Decrease in SWV
S1FcTT$_{RR}$	-8	88
S1FcCC$_{RR}$	-6	66
S1FcAA$_{SS}$	-4	94
S1FcGG$_{SS}$	TBD	TBD

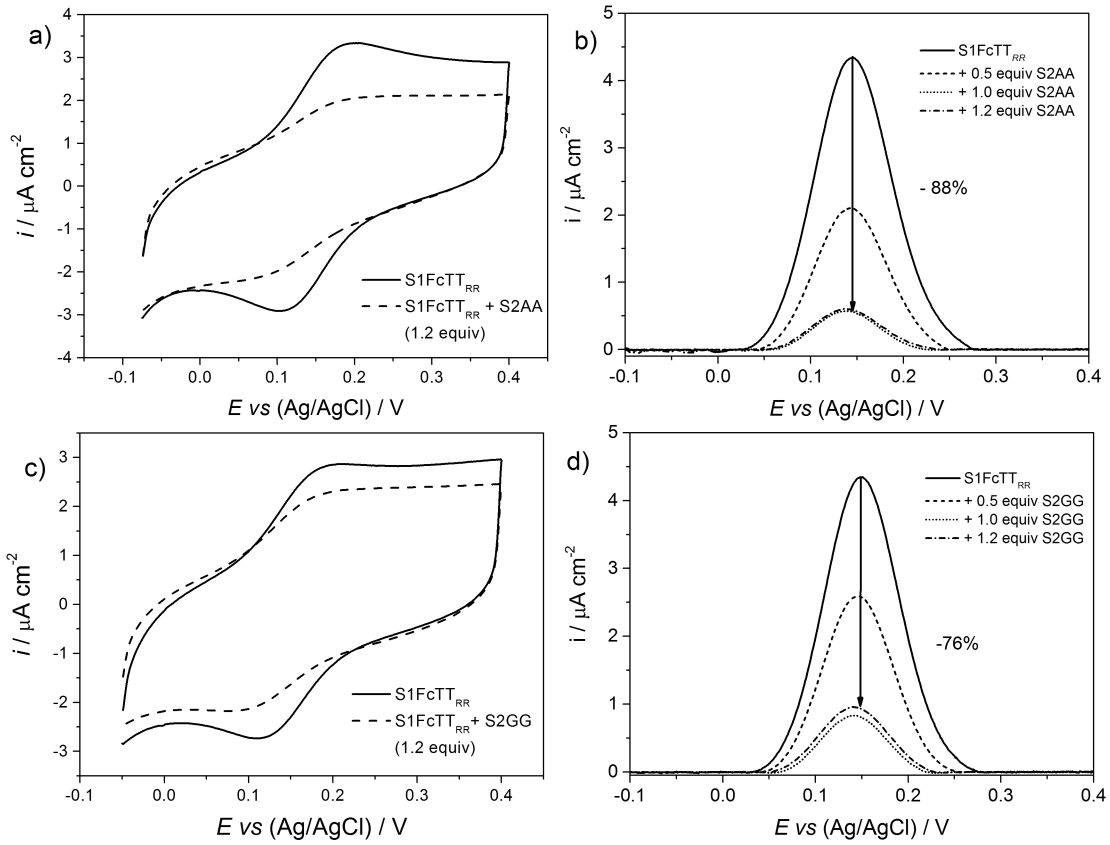


*Figure 3.44 - CV's (left) and SWV's (right) of **S1FcTT$_{RR}$** upon addition of a,b) complementary (**S2AA**) and c,d) double mismatch (**S2GG**) targets. 50 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$.*

The same signal suppression behaviour was observed upon addition of complementary target to the cytosine (-66%) and adenine (-94%) systems (Figure 3.45), however due to

depletion of oligonucleotide stocks the double mismatch targets were not able to be tested, and therefore remains an interest for future work. Furthermore, additional repeats are needed to fully evaluate the electrochemical sensing behaviour of these systems. Nevertheless, any base discriminating behaviour would also need to be observed with surface immobilised systems on electrodes if these FcNA modified oligonucleotides are to be used in biosensing applications.

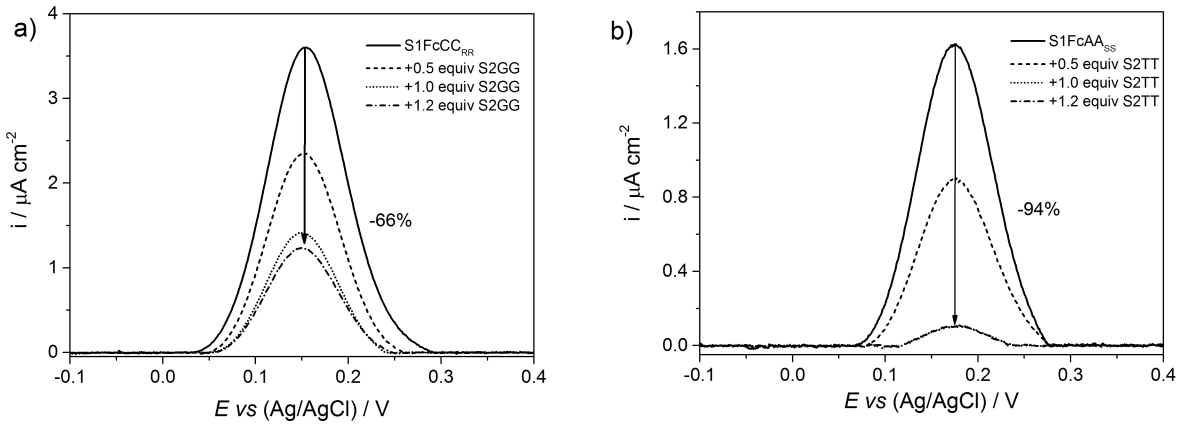


*Figure 3.45 - SWV of a) **S1FcCC$_{RR}$** upon addition of complementary target **S2GG**, and b) **S1FcAA$_{SS}$** upon addition of complementary target **S2TT**. 50 µM oligonucleotide. 50 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$.*

3.2.8 Face-to-Face FcNA Systems

It was of interest to investigate whether FcNA units could partake in complementary base pair formation with FcNA units in target strands (Figure 3.46). Therefore supported by flanking DNA base pairs, the pyrimidine and purine FcNA monomers in **S1** and **S2** respectively were evaluated for their duplex stability (Table 3.17).

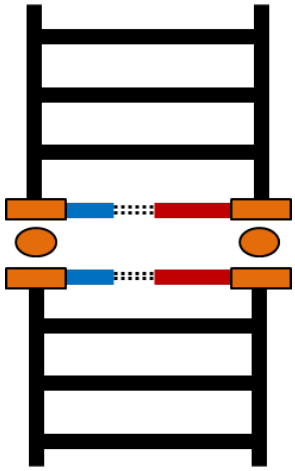


Figure 3.46 - A schematic representation of face-to-face FcNA units incorporated into DNA.

The highest observed T_m value (49.5 °C) for FcNA modified DNA oligonucleotides was observed when both incorporations had ***SS*** (***pRpR*** stereochemistry), which represents only 5.5°C lower than the value for the unmodified duplex **S1TT•S2AA** (Table 3.8**,** 56 °C). The T_m might have been a result of optimal stacking interactions in a 'zipper-like' arrangement[102,108] and therefore a double mismatch duplex with **FcCC$_{RR}$** as a control was compared with the thymine equivalent which resulted in a decrease in the overall T_m, which suggests the duplex is supported by hydrogen-bond interactions.

*Table 3.17 - Tm values for face-to-face FcNA incorporations with the FcNA pyrimidine in **S1** and **FcAA$_{RR/SS}$** in **S2**. 5 µM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	Thermal Melting Temperature (T_m) / °C	
S1 XX =	**S2FcAA$_{RR}$**	**S2FcAA$_{SS}$**
FcTT$_{RR}$	46.0 (±0.0)	46.5 (±0.3)
FcTT$_{SS}$	46.0 (±0.0)	49.5 (±0.2)
FcCC$_{RR}$	44.5 (±0.3)	TBD
FcCC$_{SS}$	TBD	TBD
FcAA$_{RR}$	45.5 (±0.0)	45.0 (±0.3)
FcAA$_{SS}$	TBD	TBD

*Highlighted values represent 'fully complementary' sequences. (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

3.2.9 FcNA Octamers

As previously discussed, synthesised octamers of **FcTT$_{RR}$** ((**FcTT$_{RR}$**)$_8$) and **FcTT$_{SS}$** ((**FcTT$_{SS}$**)$_8$) (where the number of carbons to the base = 2 and number of carbons to phosphate = 1) showed no ability to form stable duplexes with their unmodified DNA counterpart (**A$_{16}$**).[89] Furthermore an octamer bearing adenines ((**FcAA$_{SS}$**)$_8$) instead of thymines showed similar behaviour with its complement (**T$_{16}$**). This is perhaps not surprising since polymeric nucleobase systems such as these can adopt unusual conformational structures and can also suffer from solubility problems. Nevertheless these octamers were duplexed with PNA versions and the enhanced stability offered by DNA•PNA interactions due to the lack of repulsive electrostatics made it an interesting target. The 16-mer complementary strands showed a melting curve transition for the thymine derivatives (Table 3.18). The stabilities were also higher in the absence of salt in the buffer solution. The removal of salt from these solutions has been previously shown to aid duplex stability in DNA•PNA duplexes.[109] As expected no transition was observed for the adenine derivative. Similar results have been observed previously for DNA•PNA duplexes of polymeric nucleobases, where it was shown to be important that the pyrimidine was in the DNA (charged) strand.[110] When the purine was in the DNA strand, no T_m was observed; it was thought that this was due to a significant structural barrier that results in large increases in enthalpy.

Table 3.18 - T_m values for FcNA octamers duplexed with PNA targets. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$. (A.S data)

	T_m / °C	
Strand	PNA (**A$_{16}$P/T$_{16}$P**)	PNA without salt (**A$_{16}$/T$_{16}$**)
(FcTT$_{RR}$)$_8$	30.0 (**A$_{16}$P**)	35.0 (**A$_{16}$P**)
(FcTT$_{SS}$)$_8$	29.0 (**A$_{16}$P**)	33.0 (**A$_{16}$P**)
(FcAA$_{SS}$)$_8$	No T_m (**T$_{16}$P**)	No T_m (**T$_{16}$P**)
(FcHH$_{RR}$)$_8$	No T_m (**T$_{16}$P**)	No T_m (**T$_{16}$P**)

Error in instrument data = ±0.5 °C.

CD analysis of these duplexes gave better evidence for an interaction between the FcNA and PNA oligomers. The single stranded FcNA octamers display bands that are mirror images of each other with the ***SS*** isomer (blue line) showing negative bands at 450 nm and 265 nm for the d-d transitions of ferrocene and the appended nucleobases respectively (Figure 3.47a). Both bands are positive for the ***RR*** isomer (red line). These findings are in contrast to those seen for the monomers alone (Figure 3.19) which show the d-d band and π-π* transitions being the opposite sign of each other. The PNA derivative of $\mathbf{A_{16}}$ ($\mathbf{A_{16}P}$ – black line) has a slight negative band at 255 nm. PNA has no chirality in its backbone, but the observed band is likely to be a result of staggered nucleobase stacking interactions causing a slight helicity in the single strand. When duplexed with the two **FcTT** octamers, there is a dramatic change in the CD spectra (Figure 3.47b). Presumably driven by the chirality of the FcNA octamer, a large band appears at 245 nm; for the ***SS*** isomer this is negative and positive for the ***RR*** isomer. The d-d band at 450 nm remains unchanged in both sign and intensity. The region between 250 nm and 300 nm is made more complicated by the helicity of the PNA strands.

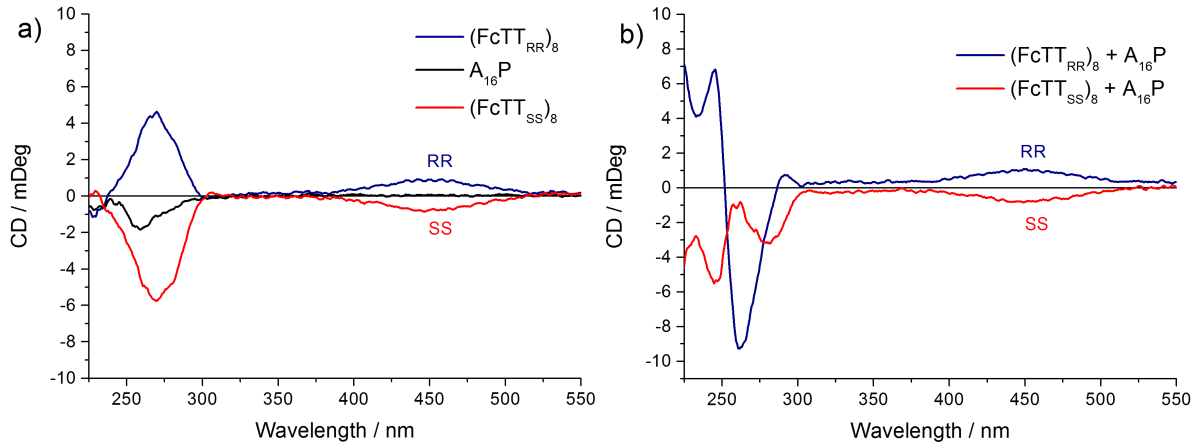


Figure 3.47 - CD spectra for the single strands of FcNA octamers ***(FcTT)$_8$*** *and* ***A$_{16}$P*** *and b) the spectra of the duplexes of* ***(FcTT)$_8$*** *and* ***A$_{16}$P****. 5 μM of each strand in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaClO$_4$.*

3.3 Conclusions

Anchoring FcNA monomers into the backbone of DNA oligonucleotides has provided a platform for comparing and optimising the structure of FcNA monomers with thymines attached. Varying the length, side chains and stereochemistry of the nucleobase and phosphate linkers results in subtle differences in overall duplex stability. MD simulations carried out by Dr. Jean-Louis Duprey support the observed stabilities. The importance of planar chirality over central chirality is evident in the observed duplex stabilities, however the methyl group in the α position does contribute towards the overall binding strength. The monomer design **Fcn3c2BB$_{XXpYpY}$** was used for the cytosine and adenine strands and the duplex stabilities were significantly enhanced compared with their thymine counterparts, these results can be explained by increased numbers of hydrogen bonds and stacking interactions for the cytosine and adenine systems respectively. Importantly for potential sensing applications, these systems exhibit large differences in base pairing affinities and hence discrimination behaviour between matching and mismatching targets. CD spectroscopy analysis of these systems supports the observed behaviour and provides insight into the structural changes afforded by these incorporations.

All FcNA modified oligonucleotides display *quasi*-reversible electrochemical behaviour due to the Fc/Fc^+ redox couple, with the $E_{1/2}$ value of each system dependent on the nucleobases appended to the Cp rings. Initial studies of the electrochemical sensing behaviour (in solution) of the systems suggests subtle differences for the thymine derivative **S1FcTT$_{RR}$** in line with differences in base pair affinities. Further investigation of this sensing behaviour with the cytosine, adenine and guanine systems would be required to fully rationalise these trends.

Finally, these FcNA-DNA conjugates have shown the ability to form H-bonded base pairs when opposite their complementary FcNA counterparts in the middle of a DNA sequence.

These systems displayed the highest observed base pair affinities to date for FcNA-DNA conjugates and remain an interesting area for future investigations.

3.4 Future Work

Whilst a number of different FcNA designs have been investigated, there remains scope for further optimising the monomer structure. Reducing the nucleobase and phosphate linker lengths, whilst maintaining the control over central and planar chirality could provide further insight into the 'ideal' FcNA monomer structure. However the trends suggest that reductions in the linker length would result in reduced duplex stabilities owing to a reduction in monomer flexibility once incorporated into DNA.

It would also be of great interest to complete the full canonical alphabet for the optimal Fc monomer design and once synthesised, it is hoped that guanine FcNA monomers would provide the largest duplex stabilities and most enhanced base pairing affinities as guanine encapsulates the characteristics of both cytosine (3 H-bonds) and adenine derivatives (purine stacking). For FcNA incorporations into DNA it would be important to carry out the following studies:

- A full T_m and CD analysis of FcGG for both stereoisomers, varying the bases opposite the monomer.
- A complete investigation by molecular dynamic simulations of all permutations of FcNA monomers (i.e. effect of appended nucleobase – A, G, C, T and effect of the bases opposite)
- A comparison of the duplex stabilities of face-to-face FcNA monomers for all four FcNA monomers (A, G, C, T) and for both ***RR*** and ***SS*** stereoisomers to gain a greater understanding of the optimal conditions for generating full FcNA-FcNA duplexes.

Following the completion of these studies, it would be exciting to build oligonucleotides containing multiple FcNA monomers with different appended nucleobases in order to build complementary FcNA duplexes.

To fully evaluate the biosensing potential of these FcNA modified oligonucleotides, it is recommended that a full study of all four FcNA monomers in strand **S1** bound to different targets in solution is completed. More importantly, these systems should be appended to a surface which is vital for the design of rapid response, re-usable biosensing devices. Being able to distinguish between single nucleotide changes in DNA sequences is of interest for the development of personalised diagnostic devices.

3.5 References

1. P. Hunter, *EMBO Rep.*, 2013, **14**, 410-413.
2. V. B. Pinheiro and P. Holliger, *Curr. Opin. Chem. Biol.*, 2012, **16**, 245-252.
3. A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh and J. Wengel, *J. Am. Chem. Soc.*, 1998, **120**, 13252-13253.
4. A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607-3630.
5. A. Eschenmoser, *Science*, 1999, **284**, 2118-2124.
6. K. U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347-1351.
7. J. K. Ichida, A. Horhota, K. Zou, L. W. McLaughlin and J. W. Szostak, *Nucleic Acids Res.*, 2005, **33**, 5219-5225.
8. C. Hendrix, H. Rosemeyer, I. Verheggen, A. Van Aerschot, F. Seela and P. Herdewijn, *Chem. Eur. J*, 1997, **3**, 110-120.
9. K. Vastmans, S. Pochet, A. Peys, L. Kerremans, A. Van Aerschot, C. Hendrix, P. Marlière and P. Herdewijn, *Biochemistry*, 2000, **39**, 12757-12765.
10. V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S.-Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn and P. Holliger, *Science*, 2012, **336**, 341-344.
11. P. Herdewijn and E. De Clercq, *Biorg. Med. Chem. Lett.*, 2001, **11**, 1591-1597.
12. J. Wang, B. Verbeure, I. Luyten, M. Froeyen, C. Hendrix, H. Rosemeyer, F. Seela, A. Van Aerschot and P. Herdewijn, *Nucleosides Nucleotides Nucl. Acids*, 2001, **20**, 785-788.
13. L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174-4175.
14. J. Summerton and D. Weller, *Antisense Nucleic Acid Drug Dev.*, 1997, **7**, 187-195.
15. N. Langkjær, A. Pasternak and J. Wengel, *Biorg. Med. Chem.*, 2009, **17**, 5420-5425.
16. V. Rait, D. Sergueev, J. Summers, K. He, F. Huang, B. Krzyzanowska and B. R. Shaw, *Nucleosides Nucleotides*, 1999, **18**, 1379-1380.
17. H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard and E. T. Kool, *Science*, 2003, **302**, 868-871.
18. H. Liu, J. Gao and E. T. Kool, *J. Org. Chem.*, 2005, **70**, 639-647.
19. S. Spitzer and F. Eckstein, *Nucleic Acids Res.*, 1988, **16**, 11691-11704.
20. R. Cosstick and J. S. Vyle, *J. Chem. Soc., Chem. Commun.*, 1988, 992-993.
21. R. Cosstick and J. S. Vyle, *Tetrahedron Lett.*, 1989, **30**, 4693-4696.

22. J. S. Vyle, B. A. Connolly, D. Kemp and R. Cosstick, *Biochemistry*, 1992, **31**, 3012-3018.
23. H. Brisset, A. E. Navarro, N. Spinelli, C. Chaix and B. Mandrand, *Biotechnol J*, 2006, **1**, 95-98.
24. T. Ihara, M. Nakayama, M. Murata, K. Nakano and M. Maeda, *Chem. Commun.*, 1997, 1609-1610.
25. M. Nakayama, T. Ihara, K. Nakano and M. Maeda, *Talanta*, 2002, **56**, 857-866.
26. J. Qiu, A. H. El-Sagheer and T. Brown, *Chem. Commun.*, 2013, **49**, 6959-6961.
27. A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388-1405.
28. P. Nielsen, M. Egholm, R. Berg and O. Buchardt, *Science*, 1991, **254**, 1497-1500.
29. M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Nordén and P. E. Nielsen, *Nature*, 1993, **365**, 566-568.
30. X. Luo, T. M.-H. Lee and I.-M. Hsing, *Anal. Chem.*, 2008, **80**, 7341-7346.
31. N. Hüsken, M. Gębala, A. Battistel, F. La Mantia, W. Schuhmann and N. Metzler-Nolte, *Chemphyschem*, 2012, **13**, 131-139.
32. N. Hüsken, M. Gębala, W. Schuhmann and N. Metzler-Nolte, *ChemBioChem*, 2010, **11**, 1754-1761.
33. E. Socher, A. Knoll and O. Seitz, *Org. Biomol. Chem.*, 2012, **10**, 7363-7371.
34. P. E. Nielsen and G. Haaima, *Chem. Soc. Rev.*, 1997, **26**, 73-78.
35. J. Wang and J. Zadeii, *Talanta*, 1987, **34**, 909-914.
36. M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina and B. F. Vanyushin, *Nature*, 1977, **270**, 369-370.
37. I. Y. Khudyakov, M. D. Kirnos, N. I. Alexandrushkina and B. F. Vanyushin, *Virology*, 1978, **88**, 8-18.
38. G. H. Clever, C. Kaul and T. Carell, *Angew. Chem. Int. Ed.*, 2007, **46**, 6226-6236.
39. K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shiro and M. Shionoya, *J. Am. Chem. Soc.*, 2002, **124**, 12494-12498.
40. G. H. Clever, K. Polborn and T. Carell, *Angew. Chem. Int. Ed.*, 2005, **44**, 7204-7208.
41. K. Tanaka, Y. Yamada and M. Shionoya, *J. Am. Chem. Soc.*, 2002, **124**, 8802-8803.
42. D. W. Gruenwedel and M. K. Cruikshank, *Nucleic Acids Res.*, 1989, **17**, 9075-9086.
43. M. Sundaralingam and J. A. Carrabine, *Biochemistry*, 1971, **10**, 292-299.
44. J. Kondo, T. Yamada, C. Hirose, I. Okamoto, Y. Tanaka and A. Ono, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 2385-2388.
45. T. J. Kealy and P. L. Pauson, *Nature*, 1951, **168**, 1039-1040.

46. G. Wilkinson, M. Rosenblum, M. C. Whiting and R. B. Woodward, *J. Am. Chem. Soc.*, 1952, **74**, 2125-2126.
47. J. G. Osteryoung and R. A. Osteryoung, *Anal. Chem.*, 1985, **57**, 101-110.
48. A. E. G. Cass, G. Davis, G. D. Francis, H. A. O. Hill, W. J. Aston, I. J. Higgins, E. V. Plotkin, L. D. L. Scott and A. P. F. Turner, *Anal. Chem.*, 1984, **56**, 667-671.
49. H. V. Nguyen, A. Sallustrau, J. Balzarini, M. R. Bedford, J. C. Eden, N. Georgousi, N. J. Hodges, J. Kedge, Y. Mehellou, C. Tselepis and J. H. R. Tucker, *J. Med. Chem.*, 2014, **57**, 5817-5822.
50. C. Ornelas, *New J. Chem.*, 2011, **35**, 1973-1985.
51. J. A. V. Butler, *J. Chem. Soc., Faraday Trans.*, 1924, **19**, 729-733.
52. A. Mulas, Y. Willener, J. Carr-Smith, K. M. Joly, L. Male, C. J. Moody, S. L. Horswell, H. V. Nguyen and J. H. R. Tucker, *Dalton Trans.*, 2015, **44**, 7268-7275.
53. Y. Willener, K. M. Joly, C. J. Moody and J. H. R. Tucker, *J. Org. Chem.*, 2008, **73**, 1225-1233.
54. D. R. Van Staveren and N. Metzler-Nolte, *Chem. Rev.*, 2004, **104**, 5931-5985.
55. T. S. Zatsepin, S. Y. Andreev and T. Hianik, *Russ. Chem. Rev.*, 2003, **72**, 537-554.
56. J. Duprey and J. Tucker, *Chem. Lett.*, 2014, **43**, 157-163.
57. H. Cai, X. Cao, Y. Jiang, P. He and Y. Fang, *Anal. Bioanal. Chem.*, 2003, **375**, 287-293.
58. C. E. Immoos, S. J. Lee and M. W. Grinstaff, *ChemBioChem*, 2004, **5**, 1100-1103.
59. A. Anne, B. Blanc and J. Moiroux, *Bioconj. Chem.*, 2001, **12**, 396-405.
60. I. Willner and M. Zayats, *Angew. Chem. Int. Ed.*, 2007, **46**, 6408-6418.
61. Y. Xiao, A. A. Lubin, A. J. Heeger and K. W. Plaxco, *Angew. Chem. Int. Ed.*, 2005, **44**, 5456-5459.
62. T. Ihara, Y. Maruo, S. Takenaka and M. Takagi, *Nucleic Acids Res.*, 1996, **24**, 4273-4280.
63. A. Anne, A. Bouchardon and J. Moiroux, *J. Am. Chem. Soc.*, 2003, **125**, 1112-1113.
64. C. J. Yu, Y. Wan, H. Yowanto, J. Li, C. Tao, M. D. James, C. L. Tan, G. F. Blackburn and T. J. Meade, *J. Am. Chem. Soc.*, 2001, **123**, 11155-11161.
65. F. Patolsky, Y. Weizmann and I. Willner, *J. Am. Chem. Soc.*, 2002, **124**, 770-772.
66. S. Brazill, N. E. Hebert and W. G. Kuhr, *Electrophoresis*, 2003, **24**, 2749-2787.
67. Y.-T. Long, C.-Z. Li, T. C. Sutherland, M. Chahma, J. S. Lee and H.-B. Kraatz, *J. Am. Chem. Soc.*, 2003, **125**, 8724-8725.
68. M. Inouye, R. Ikeda, M. Takase, T. Tsuri and J. Chiba, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 11606-11610.

69. R. Ikeda, S. Kobayashi, J. Chiba and M. Inouye, *Chem. Eur. J*, 2009, **15**, 4822-4828.
70. R. Ikeda, S. Kitagawa, J. Chiba and M. Inouye, *Chem. Eur. J*, 2009, **15**, 7048-7051.
71. J. Chiba, A. Akaishi, R. Ikeda and M. Inouye, *Chem. Commun.*, 2010, **46**, 7563-7565.
72. A. E. Navarro, N. Spinelli, C. Chaix, C. Moustrou, B. Mandrand and H. Brisset, *Biorg. Med. Chem. Lett.*, 2004, **14**, 2439-2441.
73. A. E. Navarro, *Nucleic Acids Res.*, 2004, **32**, 5310-5319.
74. G. Chatelain, H. Brisset and C. Chaix, *New J. Chem.*, 2009, **33**, 1139-1147.
75. J. E. B. Randles, *J. Chem. Soc., Faraday Trans.*, 1948, **44**, 322-327.
76. G. Chatelain, C. Chaix, H. Brisset, C. Moustrou, F. Fages and B. Mandrand, *Sens. Actuators, B*, 2008, **132**, 439-442.
77. T. Ihara, D. Sasahara, M. Shimizu and A. Jyo, *Supramol. Chem.*, 2009, **21**, 207-217.
78. D. Osella, M. Ferrali, P. Zanello, F. Laschi, M. Fontani, C. Nervi and G. Cavigiolio, *Inorg. Chim. Acta*, 2000, **306**, 42-48.
79. P. Koepf-Maier and H. Koepf, *Chem. Rev.*, 1987, **87**, 1137-1152.
80. R. F. Shago, J. C. Swarts, E. Kreft and C. E. J. Van Rensburg, *Anticancer Res.*, 2007, **27**, 3431-3433.
81. A. Houlton, C. J. Isaac, A. E. Gibson, B. R. Horrocks, W. Clegg and M. R. J. Elsegood, *J. Chem. Soc., Dalton Trans.*, 1999, 3229-3234.
82. K. Kowalski, A. Koceva-Chyła, A. Pieniążek, J. Bernasińska, J. Skiba, A. J. Rybarczyk-Pirek and Z. Jóźwiak, *J. Organomet. Chem.*, 2012, **700**, 58-68.
83. S. Trakossas, E. Coutouli-Argyropoulou and D. J. Hadjipavlou-Litina, *Tetrahedron Lett.*, 2011, **52**, 1673-1676.
84. K. Kowalski, J. Skiba, L. Oehninger, I. Ott, J. Solecka, A. Rajnisz and B. Therrien, *Organometallics*, 2013, **32**, 5766-5773.
85. P. James, J. Neudörfl, M. Eissmann, P. Jesse, A. Prokop and H.-G. Schmalz, *Org. Lett.*, 2006, **8**, 2763-2766.
86. M. Iurlo, L. Mengozzi, S. Rapino, M. Marcaccio, R. C. Perone, S. Masiero, P. Cozzi and F. Paolucci, *Organometallics*, 2014, **33**, 4986-4993.
87. A. N. Patwa, S. Gupta, R. G. Gonnade, V. A. Kumar, M. M. Bhadbhade and K. N. Ganesh, *J. Org. Chem.*, 2008, **73**, 1508-1515.
88. H. V. Nguyen, A. Sallustrau, L. Male, P. J. Thornton and J. H. R. Tucker, *Organometallics*, 2011, **30**, 5284-5290.
89. H. V. Nguyen, Z.-y. Zhao, A. Sallustrau, S. L. Horswell, L. Male, A. Mulas and J. H. R. Tucker, *Chem. Commun.*, 2012, **48**, 12165-12167.
90. L. Ramaley and M. S. Krause, *Anal. Chem.*, 1969, **41**, 1362-1365.

91. A. Sallustrau, PhD Thesis, University of Birmingham, 2013.
92. N. Moran, D. M. Bassani, J. P. Desvergne, S. Keiper, P. A. S. Lowden, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2006, 5003-5005.
93. D. Monticelli, R. Psaro, A. Pozzi, C. Dossi and S. Recchia, *Anal. Bioanal. Chem.*, 2005, **383**, 115-121.
94. J.-L. H. A. Duprey, Z.-y. Zhao, D. M. Bassani, J. Manchester, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2011, **47**, 6629.
95. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
96. J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.
97. R. Prins, A. R. Korswagen and A. G. T. G. Kortbeek, *J. Organomet. Chem.*, 1972, **39**, 335-344.
98. D. D. Popenoe, R. S. Deinhammer and M. D. Porter, *Langmuir*, 1992, **8**, 2521-2530.
99. S. W. Han, H. Seo, Y. K. Chung and K. Kim, *Langmuir*, 2000, **16**, 9493-9500.
100. D. Kang, X. Zuo, R. Yang, F. Xia, K. W. Plaxco and R. J. White, *Anal. Chem.*, 2009, **81**, 9109-9113.
101. N. Hüsken, PhD Thesis, Ruhr-University Bochum, 2011.
102. C. Brotschi and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2003, **42**, 1655-1658.
103. T. Nguyen, A. Brewer and E. Stulz, *Angew. Chem. Int. Ed.*, 2009, **48**, 1974-1977.
104. A. Mulas, PhD Thesis, University of Birmingham, 2013.
105. A. Ševčík, *Collect. Czech. Chem. Commun.*, 1948, **13**, 349-377.
106. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, Wiley, 2000.
107. A. K. Dutta, T. Sengupta, N. Vaval and S. Pal, *Int. J. Quantum Chem*, 2015, **115**, 753-764.
108. C. Brotschi, A. Häberli and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2001, **40**, 3012-3014.
109. S. Tomac, M. Sarkar, T. Ratilainen, P. Wittung, P. E. Nielsen, B. Nordén and A. Gräslund, *J. Am. Chem. Soc.*, 1996, **118**, 5544-5552.
110. A. Sen and P. E. Nielsen, *Biophys. J.*, 2006, **90**, 1329-1337.

Chapter 4

Mercury Binding FcNA-DNA Conjugates

4.1 Introduction

Metal ion interactions with nucleic acids are essential to many biological processes. For example, metal ions play a vital role in RNA folding and ribozyme activity as well as the stabilising of DNA-DNA interactions. The ability of nucleic acids to bind metals as co-factors is also of great interest to the field of DNA nanotechnology.[1] DNA's fibrous helical structure provides an ideal scaffold for designing nanoelectronic devices, and while the conductance of electrons and positive charge occurs over short-range distances in natural process such as in DNA repair,[2,3] the ability for DNA to conduct over much longer distances without degradation is not thought to be possible. Therefore many groups have exploited the structure of DNA by modifying synthesised oligonucleotides with metal-binding ligands in place of natural nucleobases, and by doing this have been able to control both structure and function.[4] Much of this area is beyond the scope of this work but two detailed reviews have been published by pioneers of the field, Shionoya[5] and Carell.[4] However the related subject of the binding of metal ions to the natural nucleobases (A, G, C, T and U) to form metal base pairs is of interest in this project and is the subject of this chapter.

4.1.1 Thymine-Mercury-Thymine Base Pairs

Arguably the most famous metal base pair is the one discovered by Sidney Katz. In 1952 he postulated that the binding of mercury(II) ions to the negatively charged DNA phosphate backbone was a result of changes in solution viscosity upon addition of mercuric chloride to a DNA solution, an effect he attributed to a change in duplex size.[6] However, his rationale changed in 1962 after the work of Thomas *et al* showed that the UV-vis spectra of the nucleobases changed upon addition of Hg^{2+}.[7] He then correctly postulated that mercury forms a 1:2 complex with thymine, where upon addition of mercury to DNA, a slippage occurs in a duplex so that two thymine residues can come together to form the T-Hg-T metal base pair with deprotonation of the two T nucleobases. (Figure 4.1).[8]

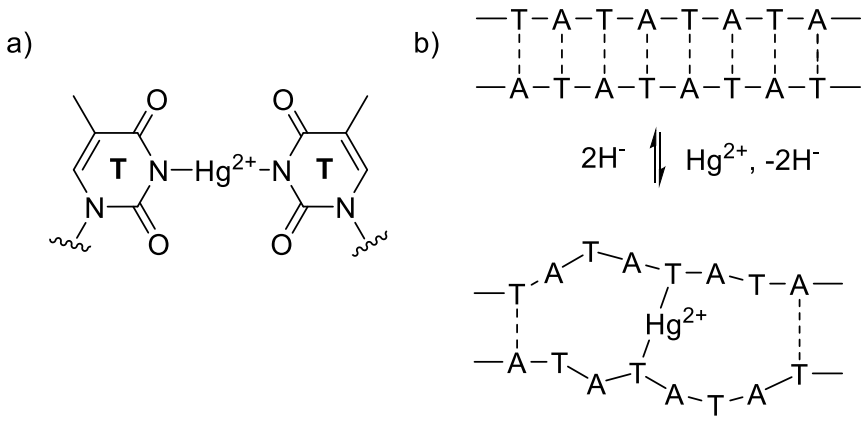


Figure 4.1 – a) The structure of T-Hg-T base pairs proposed by Katz. b) DNA 'slippage' to accommodate T-Hg-T base pairs.[8]

Since then many groups have investigated this interaction. Cruikshank *et* al found that the CD spectrum of calf-thymus DNA changed dramatically upon addition of mercury(II), indicating a large change from B-DNA to an entirely new, unknown structure.[9] Additionally spectroscopic studies have shown that the propensity of mercury being able to bind to DNA increases with the number of A-T base pairs (i.e. more T-Hg-T pairs can form). The groups of Buncel[10] and Marzilli[11] observed the emergence of new carbon and proton signals in the CH_3 and imino groups of thymine in ^{13}C and ^{1}H NMR spectra respectively when mercury was added to sequences containing T-T mismatches. They also found that the formation of more than three consecutive T-Hg-T base pairs in natural DNA was not possible, owing to intrastrand hairpin formation which was more favourable than forming a T-Hg-T stabilised duplex.[11]

More recently, the work of Ono and co-workers has been vital in confirming the exact nature of these interactions. Thermal melting studies showed that oligonucleotide duplexes containing T-T mismatches specifically bind to Hg^{2+} and upon doing so, cause a dramatic change in the overall stability of the duplex.[12] In fact for their system it was found that T-Hg-T base pairs were more stable than Watson-Crick A-T pairings and this effect was observed

across a broad pH range (3-11). Additionally electrospray mass spectrometry found that the T-Hg-T interactions were strong enough to produce molecular ion peaks. ^{15}N NMR spectroscopy showed that the $^{2}J_{NN}$ coupling constants changed upon addition of mercury to a duplex oligonucleotide containing ^{15}N(3) labelled thymines in a sequence with a double T-T mismatch.[13] Their work has also exploited this interaction to show how up to five repeating units of T-Hg-T can be arranged in a synthesised oligonucleotide duplex containing T-T mismatches. Furthermore RNA and more specifically uracil mismatches have been shown to bind Hg^{2+} as well (Figure 4.2).[14,15]

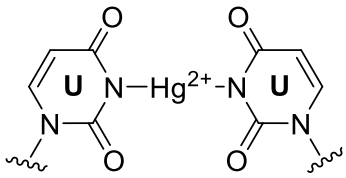


Figure 4.2 - The structure of a U-Hg-U metal base pair as described by Ono et al.[15,16]

The formation of the neutral T-Hg-T interaction is driven both enthalpically and entropically, which is respectively the result of the formation of covalent bonds and the dehydration of mercury as it becomes buried in the hydrophobic core of the DNA duplex.[17-19] In 2014, Ono and co-workers published the first x-ray crystal structure of a T-Hg-T base pair from crystals of a duplex containing a double T-T mismatch (Figure 4.3).[20]

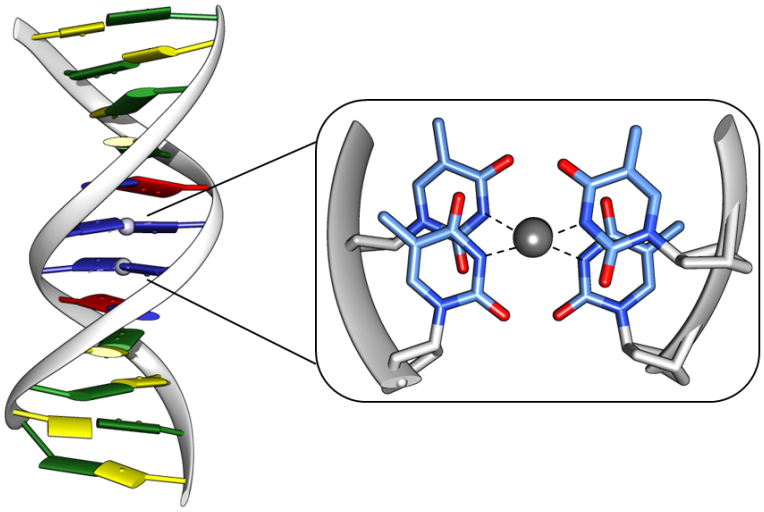


Figure 4.3 - Crystal structure of a oligonucleotide containing a double mercury mediated thymine-thymine mismatch.[20] (PDB-4L24). Visualised using Chimera II.[21]

4.1.2 Cytosine-Silver-Cytosine Base Pairs

Building on T-Hg-T systems, Ono *et al* also discovered that C-C mismatches can bind free Ag(I) as a charged adduct.[22] Again this system showed specificity for silver over other known DNA binding metals including mercury. Unlike with T-Hg-T base pairs it was found that C-Ag-C pairs only form for duplexes containing C-C mismatches and does not cause DNA slippage, which is likely a result of the increased binding strength of C-G pairs anchoring DNA in place. ^{1}H NMR investigations of a sequence containing a single C-C mismatch showed the generation of new peaks in the proton spectrum from the imino region upon addition of increasing amounts of Ag(I). A plateau at a stoichiometric 1:1 ratio of Ag:duplex again suggested a C-Ag-C pair, which was supported by both ESI-MS and thermal melting experiments.[22] Whilst it provides some indication of the possible structure, attempts to confirm it with ^{15}N NMR studies are on-going.[23]

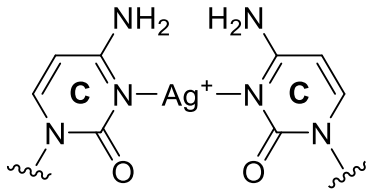


Figure 4.4 - The proposed structure of C-Ag-C base pairs in oligonucleotides containing C-C mismatches.[22]

4.1.3 Other Metal-Base Pair Complexes

Another class of metal base pair, has been described by Lee and co-workers.[24,25] At high pH, they found that divalent cations could bind to natural DNA nucleobase pairs, as evidenced by changes to the imino protons studied in ^{1}H NMR titration experiments. Modelling of the data suggested the binding of metals such as Ni^{2+}, Co^{2+} and Zn^{2+} to both A-T and G-C base pairs. The proposed structures (Figure 4.5a) involve the binding of a metal cation following deprotonation at one nucleobase; this was supported by the inability of ethidium bromide to bind the metal bound duplexes, which was ascribed to a change in charge for the duplexes. These systems have been shown to conduct like metal based wires[25-27] which is in direct contrast to unmodified/natural DNA.[2] As a result, these systems have been used in a variety of DNA arrays for electrochemical biosensing applications.[28-30] Since their discovery, the exact structure of these M-DNA interactions has been open to debate and alternative structures proposed by Lippert *et al*[31] and Alexandre *et al*[32] have been postulated (Figure 4.5b,c).

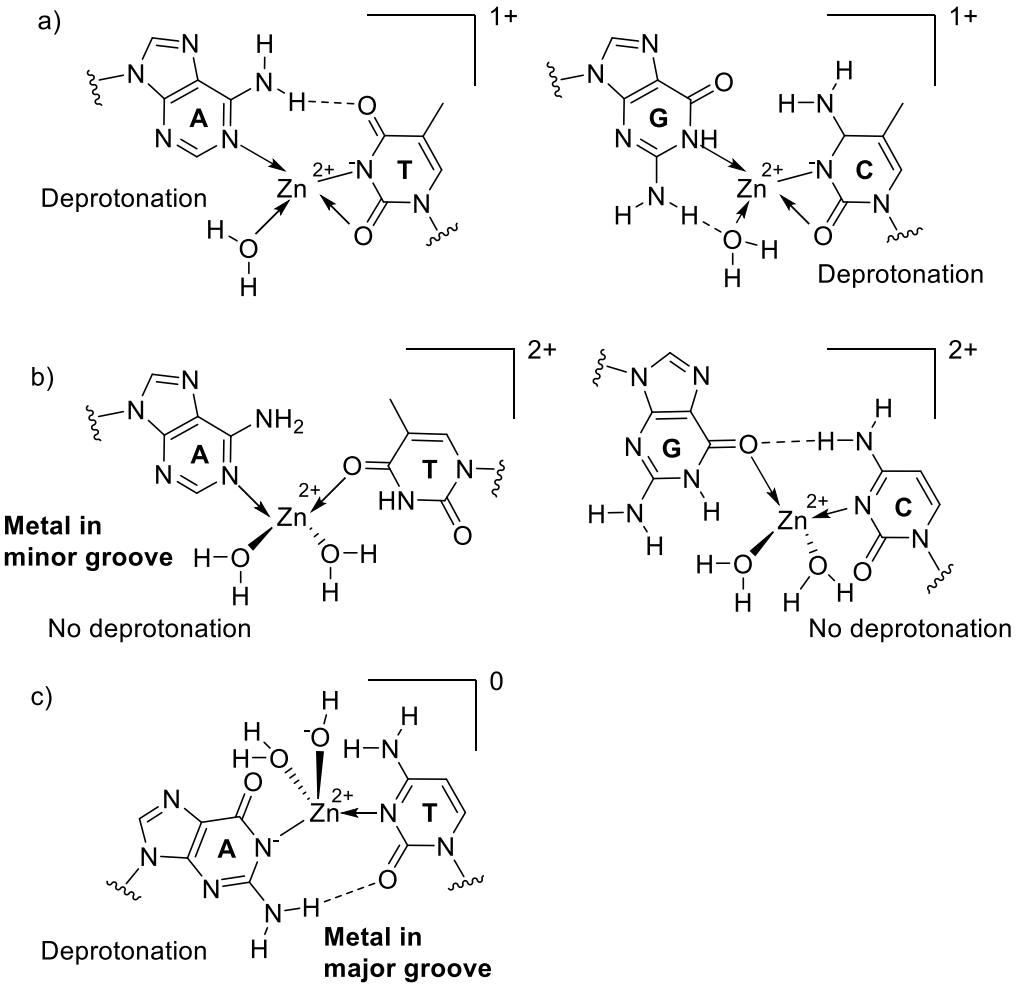


Figure 4.5 - The proposed structures of M-DNA interactions between Zn^{2+} and G-C and A-T base pairs suggested by a) Lee[25] b) Lippert[31] and c) Alexandre.[32]

4.1.4 Mercury Sensors

Accurately monitoring the levels of heavy metal pollution is a vital endeavour due to the risks posed to human health and the environment. Mercury in all forms (metallic, inorganic and organic) is known to accumulate in many bio-systems as it passes along the food chain, resulting in elevated levels in humans.[33,34] Its role in the mutation of genetic material is now well understood, with good evidence that it is brought about by DNA slippage and T-Hg-T base pair formation previously described.[8,35] Furthermore the binding of mercury to soft donor groups in major proteins and enzymes can have a detrimental effect on other important biological processes.[36] Organic mercury salts, which are formed by some aquatic

bacteria from free mercury, can cross the blood-brain barrier and have an adverse effect on many of the senses including vision, hearing and memory as well as causing complications with muscular coordination.[37,38] Inorganic mercury can cause severe damage to the vital organs, most notably the heart and kidneys.[35-37] Given the risks associated with mercury and its accumulation in both the food chain and aquatic environments, it remains important to rapidly and routinely monitor the concentration levels.[39]

Although the detection of mercury is possible by a number of established techniques such as ICP-MS[40] and absorption/emission spectroscopy,[41,42] these often require expensive equipment and labour-intensive separation protocols. Furthermore many assays involving luminescent[43-48], chromogenic[49-53] and electrochemical molecular probes,[53-55] nanoparticles[56,57] and polymers[58] have been developed. However, many of these designs have been thwarted by considerable limitations, be it selectivity, sensitivity, aqueous solubility or expense of equipment. For example organic based chromophores often only function in organic solvents and are limited by their sensitivity and response times. Excellent selectivity is a particularly important feature required for mercury sensors, as trace amounts usually coexist with large excesses of other cationic species.

Since the discovery of T-Hg-T base pairs, many oligonucleotide-based sensors have been developed in an attempt to alleviate these issues. As described the T-Hg-T interaction occurs at room temperature across a wide pH range and is highly selective for mercury over other metals, making it an ideal target for novel devices.

4.1.4.1 DNA-Nanoparticles

Many examples of oligonucleotide-based mercury sensors have utilised nanoparticle bound DNA strands that upon binding mercury, produce a colour change.[59-62] The first example published by Mirkin *et al* described a strategy based upon a mixture of two types of functionalised particle and relied upon differences in stability between T-T and T-Hg-T

containing duplexes. Each particle had conjugated to its surface one part of a complementary poly-A•poly-T duplex (Figure 4.6). In the central point of each strand was a single thymine, which upon mixing and subsequent hybridisation of the two DNA modified particles produced a T-T mismatch and aggregation of the particles. Addition of mercury to this system enhanced the duplex stability considerably. Following the addition, the samples were heated to a temperature above the T_m of the mercury free duplex, which resulted in the dissociation of these nanoparticles and a sharp colour change. However the mercury-stabilised conjugates remained aggregated and displayed a different colour.[59] Whilst assays such as this produce easily observable signal outputs, they have been limited by their sensitivity (*ca.* 1 μM).

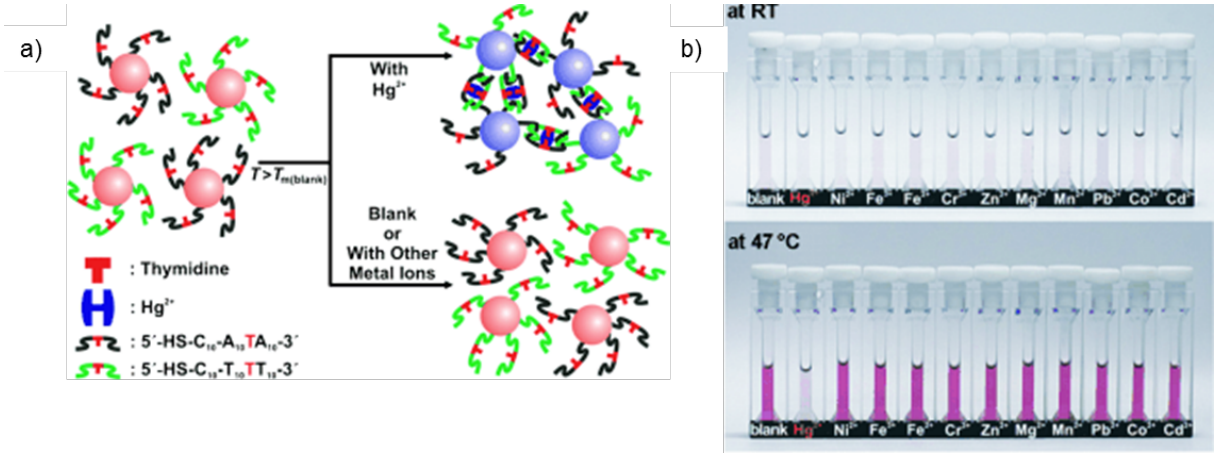


Figure 4.6 - An example of a colorimetric DNA-AuNP assay for mercury detection. a) Schematic representation of the sensing strategy. b) Colorimetric selectivity for Hg at 47 °C (1 μM mercury).[59]

4.1.4.2 Fluorescence Based Sensors

Many fluorescence based sensors for mercury rely on fluorophore-quencher interactions, in which binding mercury changes the distance between a fluorophore and a quencher. The most successful designs are based on hair-pin/stem-loop like structures similar to that shown in Figure 4.7.[63-66] Ono and co-workers described the first fluorescence-DNA sensor for mercury in 2004, which can bind up to five mercury ions, resulting in a dramatic increase in the detection limit compared with previous small molecular sensors.[53,54,67,68] The motif used

commercially available phosphoramidite moieties fluorescein and dabcyl as fluorescence and quencher respectively. As mercury binds a hair-pin like structure is formed moving the quencher closer to the fluorophore which results in increased Förster Resonance Energy Transfer (FRET) and subsequent quenching of the fluorescence signal. As with many fluorescence based DNA sensors for mercury, this design was limited in that its detection limit (40 nM) is not quite low enough for the required standards (10 nM).[65]

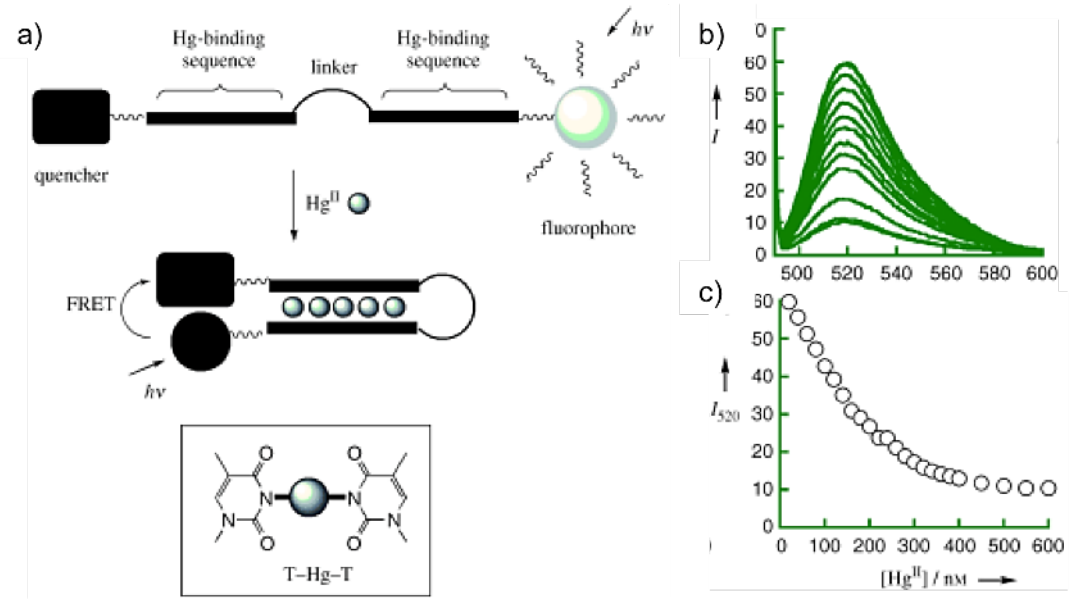


Figure 4.7 – a) Schematic representation of the hairpin design of fluorophore-quencher mercury sensor. b) The observed fluorescence output of Ono's design upon addition of Hg^{2+} (0 to 600 nM). c) Fluorescence intensity at dabcyl emission intensity (520 nm) versus the Hg^{2+} concentration.[63]

4.1.4.3 Electrochemical DNA Sensors

Surface bound DNA electrochemistry has been utilised for some time in an array of biosensing platforms for the detection of DNA,[69-72] variations in base sequences[73-75] and small molecules.[76,77] Such is the enhanced sensitivity and robustness towards false positives afforded by monolayer based designs that it perhaps the most attractive approach for the design of new DNA-based sensors. It is not surprising then that a number of examples of electrochemical assays have been developed for the detection of mercury. These include those based on impedance,[78,79] pulse voltammetry (SWV)[79-84] and cyclic voltammetry

(CV)[79,81-86]. Many rely on labelling of surface bound oligonucleotides with redox-active tags such as ferrocene or ruthenium complexes and are dependent on large conformational changes in the secondary structure of the immobilised probes. For example the work published by Yu *et al* in 2009 described a cooperative strand design in which ferrocene labelled poly-T sequences were immobilised on a surface.[79] Upon the binding of mercury to the thymine residues, a zipping up of adjacent strands on the surface occurred resulting in a more rigid duplex, consequently the ferrocene tag was not as able to move towards the electrode surface, and hence a drop in the observed oxidation current was observed, a so called 'signal-OFF' sensor (Figure 4.8). This design has a mercury detection limit of 0.5 nM and showed good selectivity but was also sensitive to the addition of Cu^{2+}.[79] This was alleviated by the addition of free adenine which complexed the Cu ions. Furthermore, its use as a detection assay was limited in that the signal-to-noise ratio was relatively small as a result of the large distance between the ferrocene tag and the electrode surface.

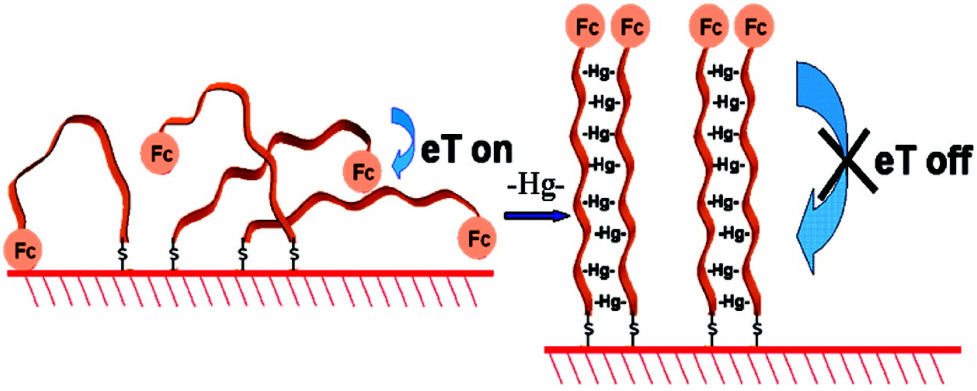


Figure 4.8 - Cooperative interstrand electrochemical sensor for mercury based on a conformational switch of ferrocene labelled poly-T sequences.[79]

More recently Yu *et al* described a structure switching design (Figure 4.9). The tagged strand itself was not directly immobilised onto the surface but instead bound as a complementary strand. It relied on a structural change-induced release of this strand upon mercury binding, which caused a self-complementary hair-pin to form. To date this is the

most sensitive oligonucleotide-based mercury sensor described, with detection limits as low as 0.06 nM.[82] This level of sensitivity is considerably lower than the toxicity threshold defined by the US Environmental Protection Agency (EPA) for drinking water (<10 nM).[87] Not only were the detection limits achieved unprecedented, but the design was also unchanged by the addition Cu^{2+} which has been problematic in other electrochemical assays.[79] This enhanced selectivity was thought to be a result of Cu^{2+} interactions being too weak to induce a stable folding of the probe strand.

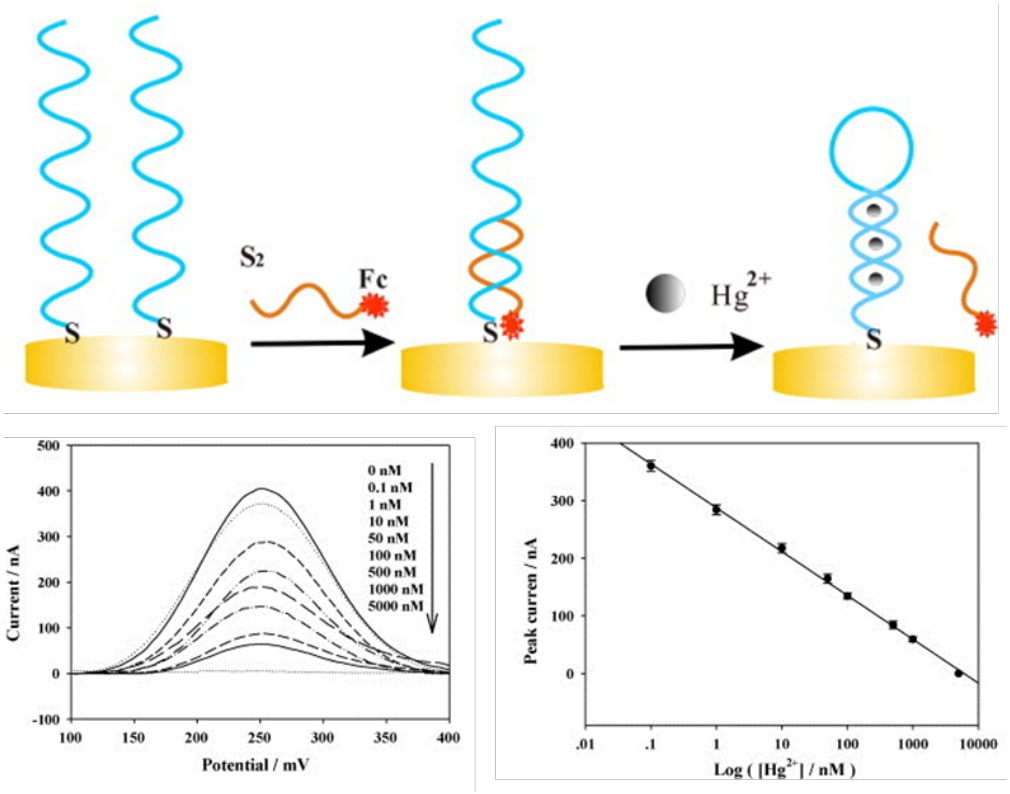


Figure 4.9 – Mercury binding induced release of a ferrocene labelled oligonucleotide causes signal suppression.[82]

4.1.5 Conclusions and Project Aims

The T-Hg-T base pair adduct is now well understood and this coupled with the health risks imposed by mercury pollution, make it an interesting target for the development of oligonucleotide based sensors. Many assays to date have been described, a majority of which utilise optical spectroscopy techniques. Electrochemical designs have become

increasingly more popular in recent years due to the enhanced sensitivity and robustness to contamination by complex sample matrices. Whilst the preparation of DNA-surfaces for electrochemistry requires optimisation for each system, reproducibility is generally considered to be excellent. Furthermore, immobilisation of probes onto a surface has been shown to be compatible with generating re-usable mercury,[81] another advantage over other approaches.

The aim of this project was to evaluate the potential of FcNA modified oligonucleotides, where the appended nucleobases are thymines, to bind mercury (Figure 4.10). FcNA modified oligonucleotides are redox-active (see Chapter 3) and could therefore provide a foundation for the development of electrochemical sensors for mercury, where the binding site very close to the redox-active centre. Such a design would be unlike other motifs, with the sensor system relying not solely on large conformational changes. It was hoped that this may provide the possibility for reducing detection limits even further beyond that which is currently possible.

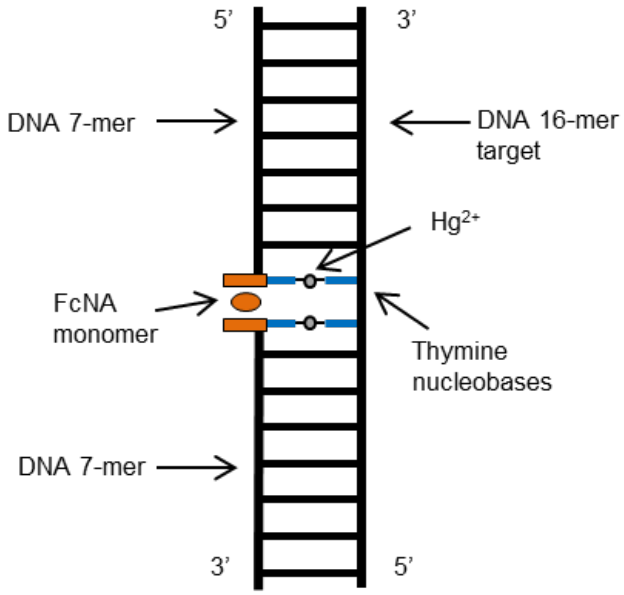


Figure 4.10 - A schematic representation of the proposed mode of Hg^{2+} binding to FcNA-DNA conjugates.

4.2 Results and Discussion

4.2.1 Thermal Melting Studies – Thymines Opposite

All of the synthesised thymine-bearing FcNA monomers (Figure 4.11) described in Chapter 3 were incorporated into **S1** and screened by thermal melting experiments in order to probe the best system for potential mercury binding.

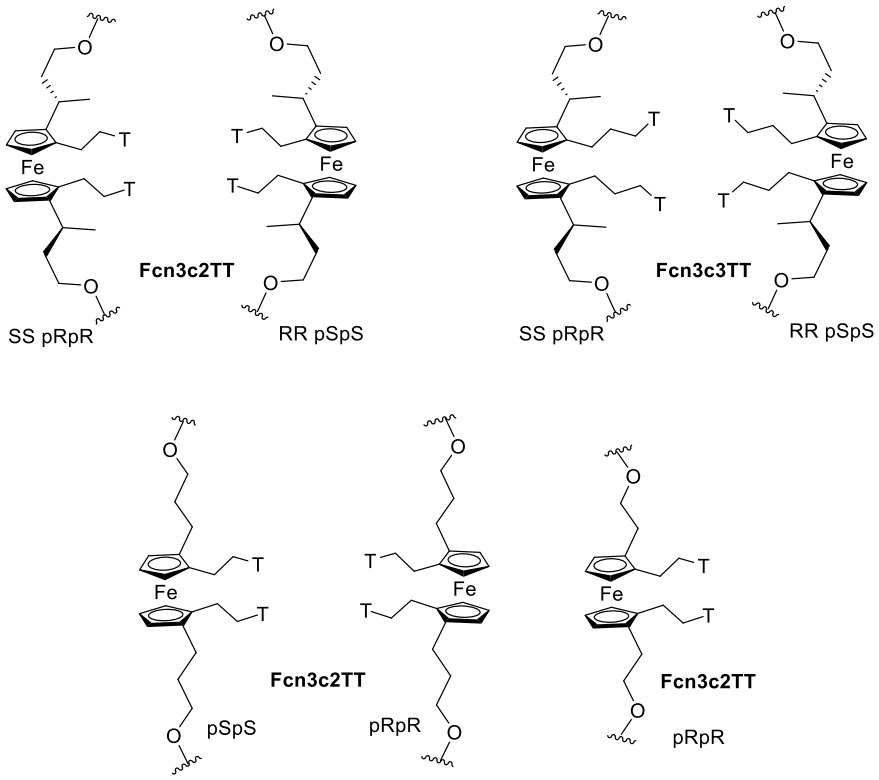


*Figure 4.11 - FcNA motifs incorporated into **S1** and **S2**.*

*Table 4.1 - Synthesised oligonucleotides **S1XX** and **S2YY** where XX = TT, FcNA monomer and YY = TT, AA or FcNA monomer).*

Strand	Sequence (5'-3')
S1XX	TGGACTC-**XX**-CTCAATG
S2YY	CATTGAG-**YY**-GAGTCCA

Each strand was hybridised with strand **S2TT** to create a double TT-TT mismatch. Table 4.2 below shows the observed T_m values of these systems in the absence and presence of 2

molar equivalents of mercury(II) perchlorate (i.e. one mercury cation for each T-T mismatch). While all but one system gave an increase in T_m, the duplex that gave the greatest stabilisation enhancement was (**S1Fcn3c2TT$_{SSpRpR}$**) with an increase of 6 °C. It is also interesting to note that the T_m value was higher than that determined when the opposite strand is its fully complementary (**S1Fcn3c2TT$_{SSpRpR}$•S2AA** = 43 °C). This suggested that binding mercury induced the most favourable conformational change, and possibly for the remaining systems the mercury was simply stabilising the negatively charged DNA backbone rather than being able to form T-Hg-T pairs. Whilst the increase was significant, the unmodified DNA equivalent produced a much larger stabilization (11.5 °C) upon the addition of mercury. This is not unexpected given the reduced flexibility of the ferrocene unit compared with a natural nucleoside pair, but also could indicate that only one T-Hg-T adduct is formed. Given that this system was significantly more promising than the others, work continued on this design alone, with the opposite stereoisomer (**S1Fcn3c2TT$_{RRpSpS}$**) used in comparison.

*Table 4.2 – T_m values for **S1XX•S2TT** upon addition of 2 molar equivalents of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C **S2TT**		
S1 XX =	**Without Hg**	**With Hg**	**Difference**
Fcn3c2TT$_{RRpSpS}$	40.5 (±0.3)	43.5 (±0.3)	+3.0
Fcn3c2TT$_{SSpRpR}$	39.5 (±0.2)	45.5 (±0.2)	+6.0
Fcn3c3TT$_{RRpSpS}$	40.5 (±0.5)	41.5 (±0.3)	+1.0
Fcn3c3TT$_{SSpRpR}$	40.5 (±0.3)	41.5 (±0.3)	+1.0
Fcn3c2TT$_{pSpS}$	40.5 (±0.5)	41.5 (±0.2)	+1.0
Fcn3c2TT$_{pRpR}$	40.5 (±0.3)	40.0 (±0.2)	-0.5
Fcn2c2TT$_{pRpR}$	41.0 (±0.1)	43.5 (±0.2)	+2.5
TT	41.5 (±0.2)	53.0 (±0.0)	+11.5

(.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

4.2.2 Thermal Melting Studies - Adenines Opposite

To confirm that the stabilisation was in fact a result of the formation of a T-Hg-T adduct, and not due to cation binding to the negatively charged phosphate backbone, the same experiments were undertaken with the control sequence **S2AA**. Significantly for all three duplexes, a reduction in duplex stability was observed (Table 4.3). However, the reduction was largest with **S1Fcn3c2TT$_{SSpRpR}$** as seen previously, with a 7 °C decrease. This could be a result of DNA slippage, in which the DNA structure rearranges to form T-Hg-T base pairs.[8] CD spectroscopy analysis of this duplex suggested that there was indeed a large structural change, as evidenced by a complete loss of the characteristic B-DNA bands (Figure 4.12).

*Table 4.3 - T_m values for **S1XX•S2AA** upon addition of 2 molar equivalents of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C S2AA		
S1 XX =	**Without Hg**	**With Hg**	**Difference**
Fcn3c2TT$_{RRpSpS}$	43.0 (±0.2)	41.5 (±0.7)	-1.5
Fcn3c2TT$_{SSpRpR}$	41.0 (±0.0)	34.0 (±0.0)	-7.0
TT	55.5 (±0.3)	54.0 (±0.0)	-1.5

(.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

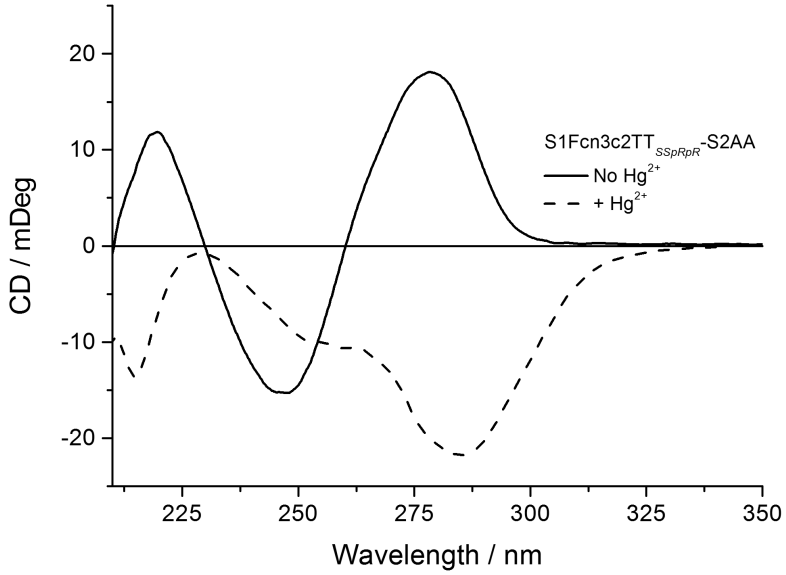


*Figure 4.12 - CD spectra for the duplex **S1Fcn3c2TT$_{SSpRpR}$ •S2AA** in the absence (solid line) and presence (dashed line) of Hg^{2+}. $Hg(ClO_4)_2$ at 2 molar equivalents. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

4.2.3 Thermal Melting Studies - Face-to-Face FcNA Thymines

Duplexes containing face-to-face FcNA thymine monomers were probed for their ability to bind mercury. In contrast to when unmodified DNA was opposite the FcNA unit in the duplex, the addition of mercury resulted in a loss of overall stability for both stereoisomers. This destabilisation is likely to be a result of the increased distortion that would be imparted by locking two FcNA units together.

*Table 4.4 - T_m values for **S1Fcn3c2TT** when duplexed with **S2YY** (where Y= the same FcNA monomer) upon addition of 2 molar equivalent of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C Opposite itself		
S1 XX =	**Without Hg**	**With Hg**	**Difference**
Fcn3c2TT$_{RRpSpS}$	44.5 (±0.3)	42.5 (±0.3)	-2.0
Fcn3c2TT$_{SSpRpR}$	45.5 (±0.3)	43.0 (±0.3)	-2.5

(.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

4.2.4 Further Studies of S1Fcn3c2TT$_{SSpRpR}$/S1FcTT$_{SS}$

The system giving the most promising, **S1Fcn3c2TT$_{SSpRpR}$** (from herein known as **S1FcTT$_{SS}$**) was investigated in more detail. All experiments throughout the rest of this chapter were completed alongside the control duplex **S1TT•S2TT** to enable comparisons between natural and FcNA modified oligonucleotides to be made.

4.2.4.1 Titration Experiments

Ono and others have previously shown that T-Hg-T causes stabilisation of T-T mismatches, which plateaus at a 1:1 stoichiometric concentration of mercury for each T-T mismatch.[12] Therefore the two systems were studied similarly by adding aliquots of mercury(II) perchlorate to duplexes and monitoring the effect by thermal melting analysis. The thermal melting curves for both **S1TT•S2TT** and **S1FcTT$_{SS}$•S2TT** in the presence of increasing amounts of mercury are shown in Figure 4.14. For both systems, the T_m value begins to plateau at 1 molar equivalent per T-T mismatch of mercury, as would be expected for a T-

Hg-T binding motif. Furthermore, it also provides evidence that both T-T mismatches are involved in binding mercury (*vide supra*). Whilst the T_m curves do continue to change as more mercury is added, the overall T_m only slightly increases, this could be an effect of driving the equilibrium further to the right or because of negative phosphate backbone stabilisation afforded by the free Hg^{2+} ions.

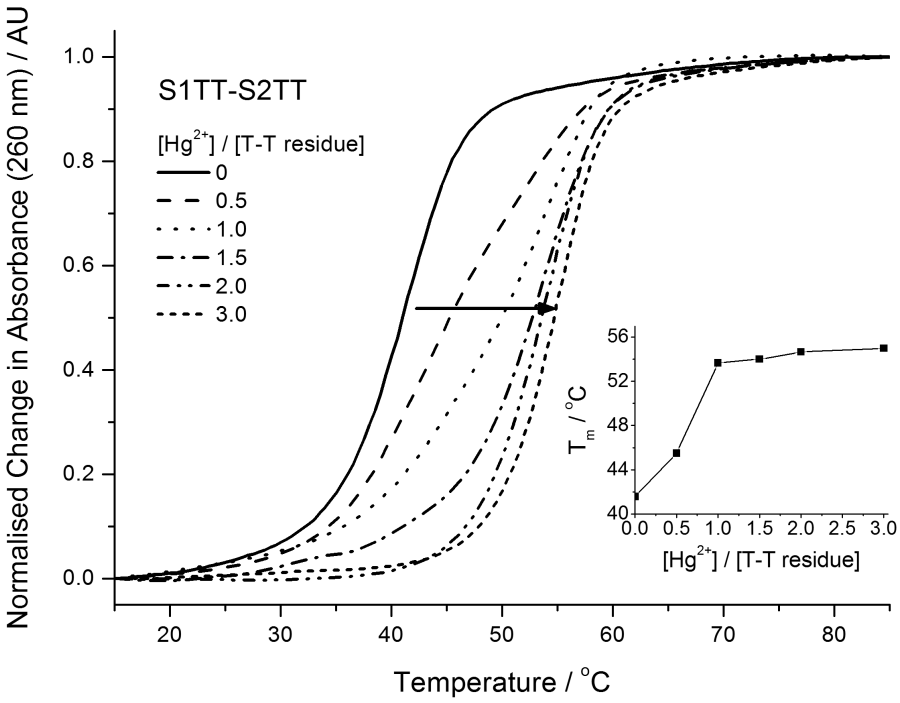


*Figure 4.13 - Thermal melting profiles of **S1TT•S2TT** with increasing amounts of mercury. (Inset) The observed change in T_m value as a function of increasing equivalents of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

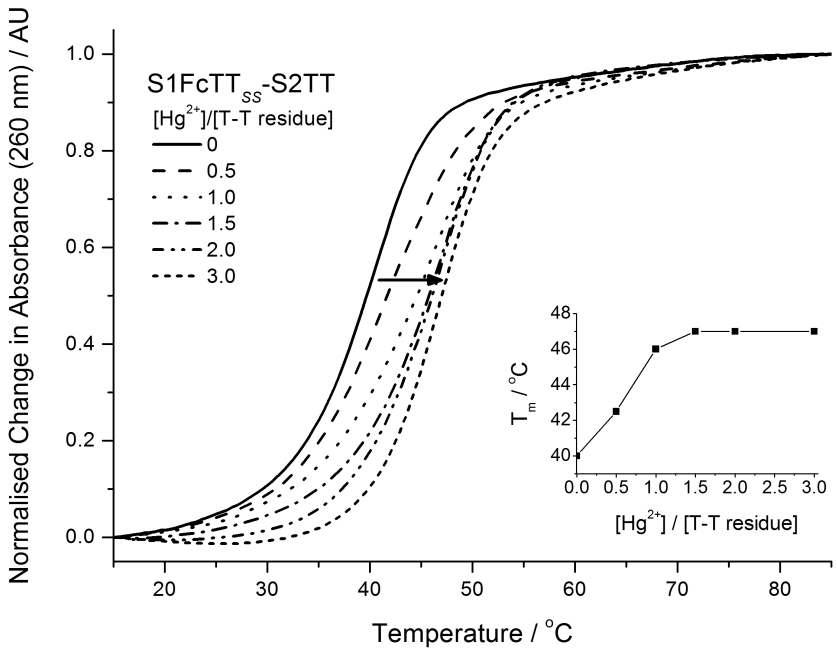


*Figure 4.14 - Thermal melting profiles of **S1FcTT$_{SS}$•S2TT** with increasing amounts of mercury. (Inset) The observed change in T_m value as a function of increasing equivalents of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

UV-vis titrations mimicking these experiments showed a red shift and reduction in the absorbance with increasing mercury concentration. The same behaviour has been noted by others[12,88] and can be explained by a tightening of the DNA duplex upon adduct formation and increase in hypochromism. Plots of the change in absorbance (at 260 nm) as a function of the mercury ion concentration (Figure 4.15) showed that for the unmodified duplex, the change levels off at one Hg ion per T-T residue, whereas for the FcNA modified sequence a slight excess is required. Control experiments with **S2AA** as the target (i.e. where mercury shouldn't bind) also showed a change upon addition of mercury (Figure 4.16, Figure 4.17). Given that the stability of these duplexes is affected by the addition of mercury, it is not surprising that some change would be observed. However, unlike with the T-T mismatch systems, the FcNA controls showed a steady shift, with a change to an increase in absorption intensity above 1.5 equivalents, indicating a large distortion in the duplex, which supports the idea that a DNA slippage occurs as proposed from the CD data in Figure 4.12.

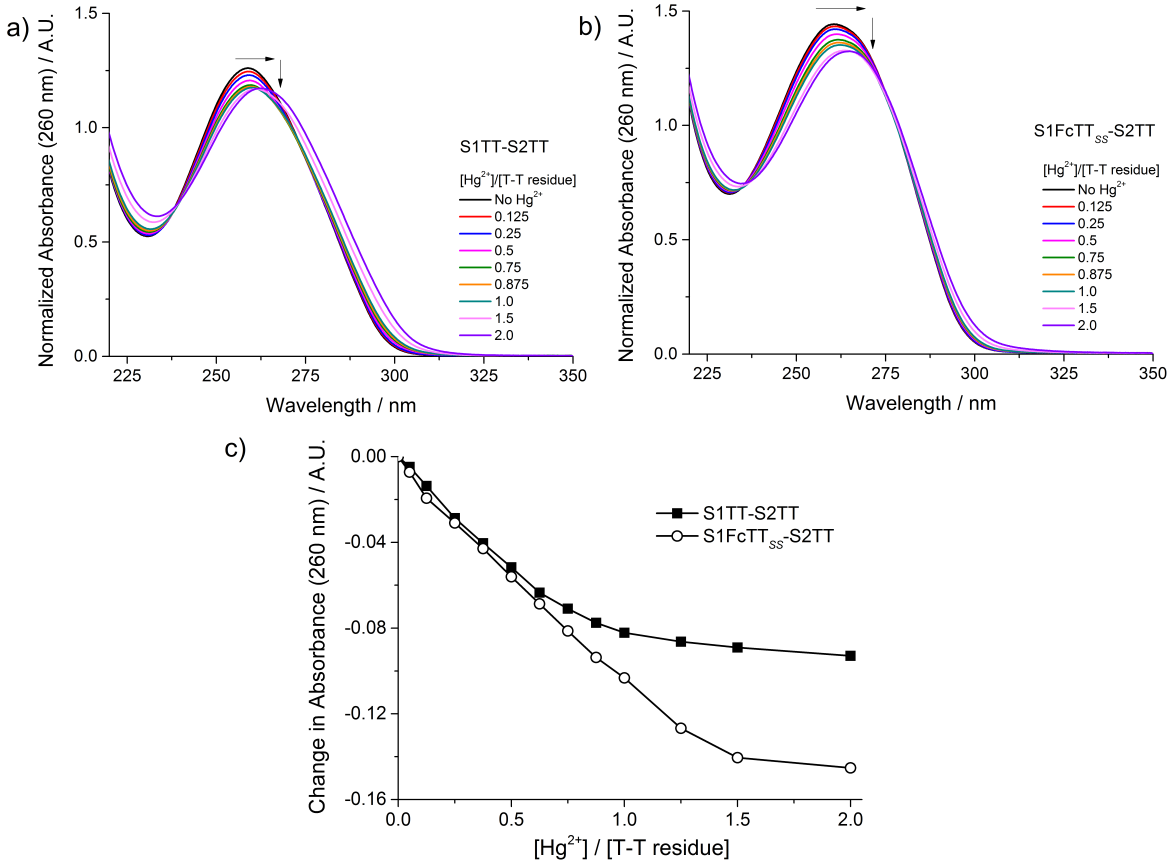


*Figure 4.15 - UV-vis spectra of a) **S1TT•S2TT** and b) **S1FcTT$_{SS}$•S2AA** upon addition of increasing amounts of $Hg(ClO_4)_2$. c) The observed change in absorbance as a function of equivalents of mercury per T-T mismatch. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

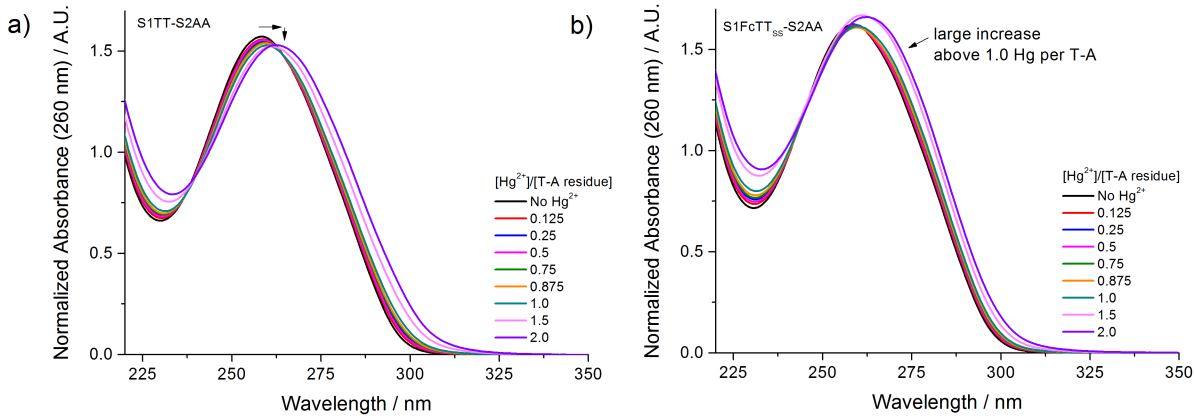


*Figure 4.16 - UV-vis spectra of a) **S1TT•S2AA** and b) **S1FcTT$_{SS}$•S2AA** upon addition of increasing amounts of $Hg(ClO_4)_2$. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

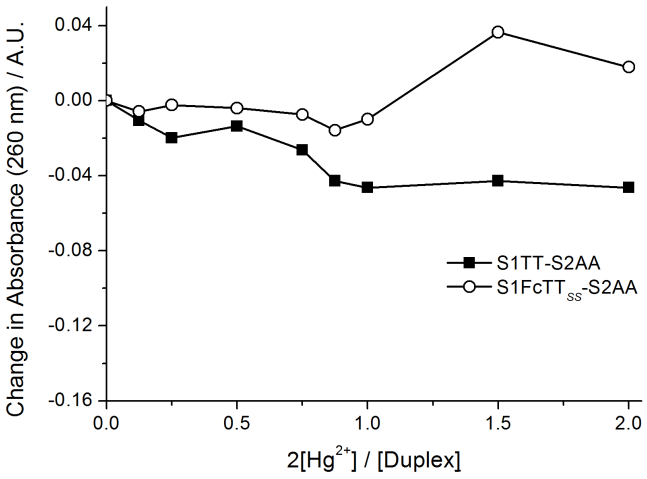


Figure 4.17 - The observed change in absorbance as a function of 2 x mercury (II) concentration with respect to duplex concentration. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.

4.2.4.2 Circular Dichroism Studies

CD spectra of both unmodified and FcNA systems in the absence and presence of mercury show dramatic differences (Figure 4.18). For both duplexes there is a red shift in both of the characteristic B-DNA bands and a large reduction in the positive band intensity, which is in line with the observed UV-vis changes (*vide supra*) This behaviour has previously been observed for shorter oligonucleotides containing T-Hg-T pairs and has been attributed to a distortion away from the classic B-DNA structure.[12,88] Whereas the band at 218 nm only differs slightly for the unmodified duplex, the band increases significantly for **S1FcTT$_{SS}$•S2TT**. This can be explained by a 'locking' of the ferrocene motif into more of a helical environment as a result of the linear T-Hg-T bonds formed with the opposing strand. This type of change has been previously observed for FcNA-DNA conjugates when bound to strands with complementary and non-complementary nucleobases opposing the ferrocene unit (Chapter 3.2.7).

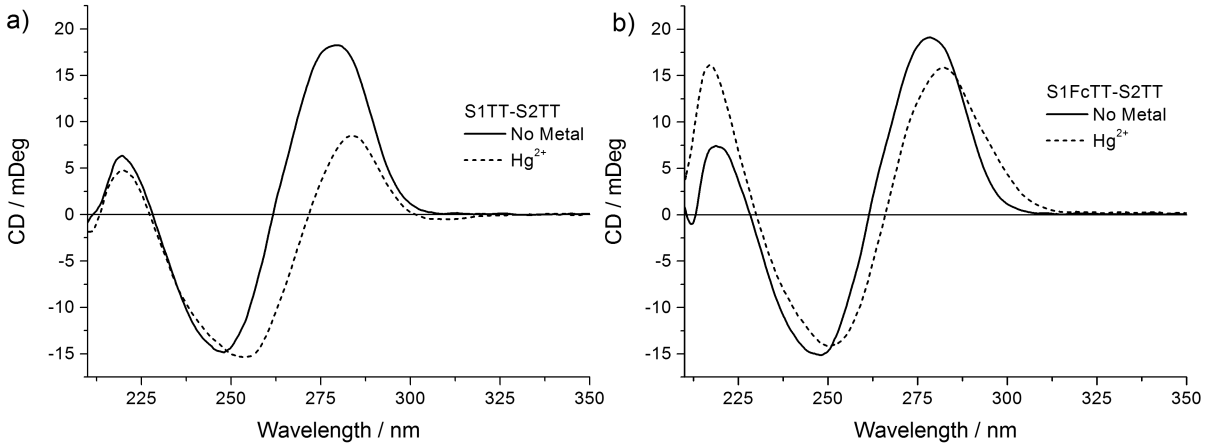


Figure 4.18 - CD spectra of a) ***S1TT•S2TT*** *and b)* ***S1Fcn3c2TT$_{SSpRpR}$•S2TT*** *in the absence (solid line) and presence of 2 molar equivalents of $Hg(ClO_4)_2$ (dashed line). 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

4.2.4.3 Selectivity for Mercury over other Metals

T-T mismatches, as discussed previously, have been shown to be highly selective towards mercury against other heavy metals and divalent cations.[12] With this in mind, it was important to investigate whether FcNA systems showed the same selectivity. Accordingly, thermal melting experiments of the duplexes with an array of different divalent cations were probed (Table 4.5). Judging by the T_m values, the selectivity for Hg^{2+} was found to be high, with relatively small increases for the other cations in both the modified and unmodified strands. The graphs in Figure 4.19 indicate the differences well, with little change to the shape of the thermal melting profiles compared to Hg systems. In fact the addition of 100-fold excess of these control metals yielded only very small further increases in the observed T_m values (Table 4.6) from those where only two molar equivalents are added. Interestingly the addition of Cu^{2+} in high excess causes a reduction in the overall stability. Large quantities of Cu^{2+} have previously been shown to reduce duplex stabilities.[89]

*Table 4.5 - T_m values for **S1FcTT$_{SS}$•S2TT** and **S1TT•S2TT** upon addition of 2 molar equivalents of other divalent cations. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C Duplex			
Metal	**S1FcTT$_{SS}$•S2TT**	ΔT_m*	**S1TT•S2TT**	ΔT_m*
No Metal	39.5 (±0.3)	-	41.5 (±0.0)	-
Hg^{2+}	46.5 (±0.2)	+7.0	53.5 (±0.0)	+12.0
Zn^{2+}	41.0 (±0.0)	+1.5	42.5 (±0.5)	+1.0
Pb^{2+}	40.5 (±0.5)	+1.0	42.0 (±0.0)	+0.5
Cd^{2+}	40.0 (±0.3)	+0.5	41.5 (±0.0)	+0.0
Fe^{2+}	40.5 (±0.3)	+1.0	42.0 (±0.3)	+0.5
Cu^{2+}	41.5 (±0.3)	+2.0	41.5 (±0.0)	+0.0
Ba^{2+}	40.0 (±0.3)	+0.5	42.0 (±0.0)	+0.5
Ca^{2+}	41.0 (±0.3)	+1.5	42.0 (±0.0)	+0.5
Mg^{2+}	40.0 (±0.3)	+0.5	42.0 (±0.0)	+0.5

* Δ*T*m is in comparison with the duplex in the absence of metal ions.

*Table 4.6 - Tm values for **S1FcTT$_{SS}$•S2TT** upon addition 100 molar equivalents of other divalent cations. Hg^{2+} at 2 molar equivalents is shown for reference. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM NaClO4.*

	T_m / °C Duplex			
Metal	**S1FcTT$_{SS}$•S2TT**	ΔT_m*	**S1TT•S2TT**	ΔT_m*
No Metal	39.5 (±0.3)	-	41.5 (±0.3)	-
Hg^{2+}	46.5 (±0.3)	+7.0	53.5 (±0.3)	+12.0
Zn^{2+}	41.5 (±0.3)	+2.0	42.5 (±0.3)	+1.0
Pb^{2+}	41.0 (±0.2)	+1.5	42.0 (±0.3)	+1.5
Cd^{2+}	40.5 (±0.0)	+1.0	42.0 (±0.3)	+0.5
Fe^{2+}	39.5 (±0.3)	+0.0	42.0 (±0.0)	+0.5
Cu^{2+}	42.5 (±0.3)	+2.0	41.0 (±0.3)	-0.5
Ba^{2+}	40.5 (±0.3)	+1.0	42.5 (±0.2)	+1.0
Ca^{2+}	41.0 (±0.3)	+1.5	42.0 (±0.2)	+0.5
Mg^{2+}	41.0 (±0.2)	+1.5	43.0 (±0.3)	+1.5

* Δ*T*m is in comparison with the duplex in the absence of metal ions.

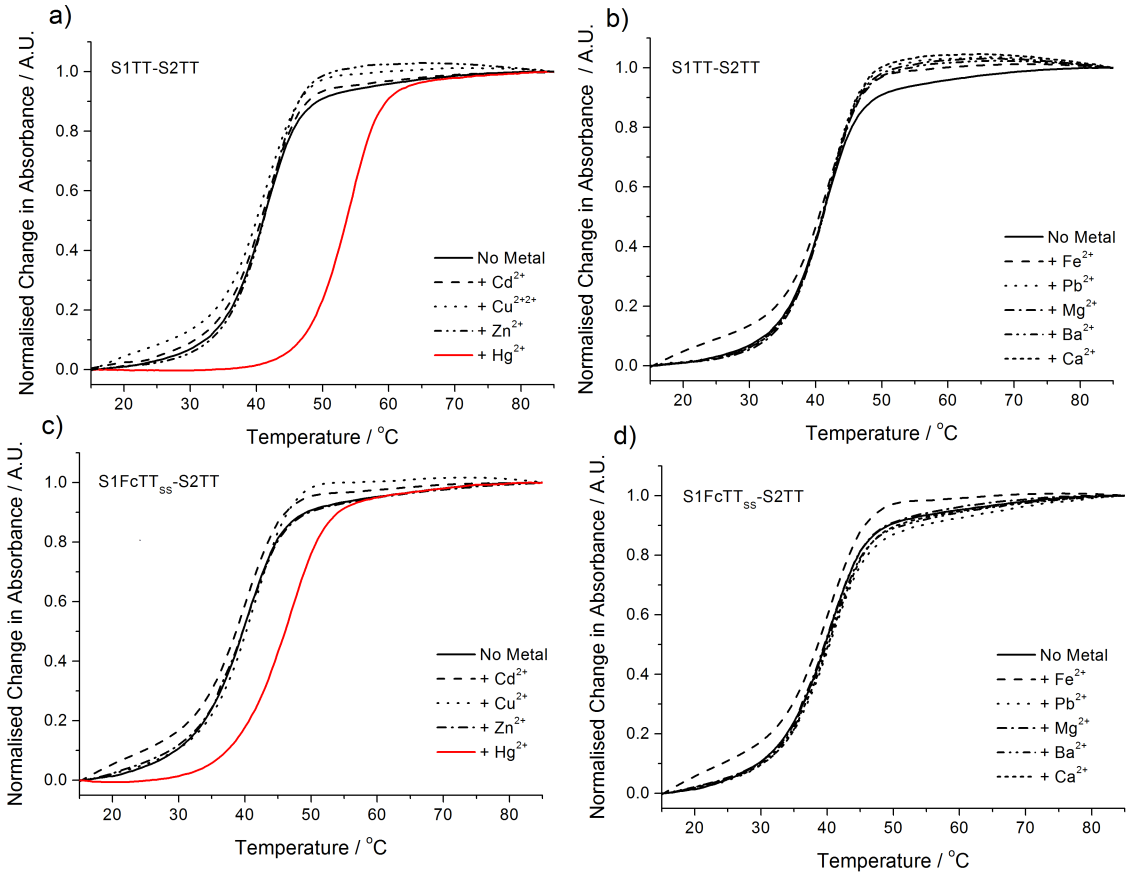


*Figure 4.19 - Addition of metals to duplexes **S1TT•S2TT** a) No metal, Cd^{2+}. $Cu^{2+,}$ $Zn^{2+,}$ Hg^{2+} b) $Fe^{2+,}$ Pb^{2+}, $Mg^{2+,}$ $Ba^{2+,}$ Ca^{2+} Fe^{2+} and **S1FcTTss•S1TT** c) No metal, Cd^{2+}. $Cu^{2+,}$ $Zn^{2+,}$ Hg^{2+} and d) $Fe^{2+,}$ Pb^{2+}, $Mg^{2+,}$ $Ba^{2+,}$ Ca^{2+} Fe^{2+}. Metals added to 2 molar equivalents. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

CD spectra of the two duplexes upon addition of a selection of other metal ions revealed that conformational changes only occurred in the presence of mercury ions. For cadmium, copper and zinc the classic B-DNA structure is retained (positive band at ~275 nm, negative band at 248 nm and positive band at 218 nm). It was interesting to note that for the FcNA modified system, a small decrease and very slight red shift in the positive band at 275 nm was observed in the presence of copper, which could be related to the slightly elevated T_m observed upon addition of this ion (Table 4.5). Copper has been shown to bind to the phosphate backbone, but also to have some binding interaction with the DNA nucleobases,[89-93] which could explain these small changes in the B-DNA bands.

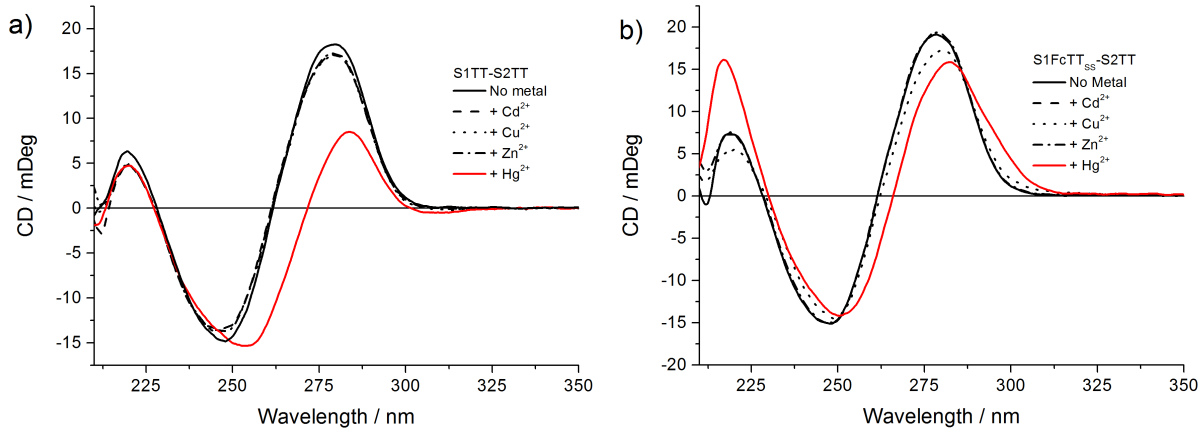


*Figure 4.20 - CD spectra of duplexes a) **S1TT•S2TT** and b) **S1FcTT$_{SS}$•S2TT** upon addition of heavy metals. No metal - solid black line, Cd^{2+} - dashed lined, Cu^{2+} - dotted line, Zn^{2+} - dash-dot line, Hg2+ - red line. Metals added to 2 molar equivalents. 5 µM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

4.2.4.4 Mass Spectrometry Characterisation

Mass spectrometry was used to support the characteristic CD and UV-vis data indicating mercury was bound to the T-T double mismatch. Both unmodified and FcNA derivatives of **S1** were bound to **S2TT** and 2 molar equivalents of mercury added to the duplexes. The resulting mercury-DNA conjugates were subjected to ionisation by electrospray ionisation. For both the unmodified DNA duplex (Figure 4.21) and the FcNA equivalent (Figure 4.22), three ions of interest were observed; the duplex, the duplex with one mercury ion bound and the duplex with two mercury ions bound. The unmodified duplex produced more detectable ions and as a result, the peaks were better defined from the background. Both samples were particularly salty despite the samples being prepared in ultra-pure Milli-Q water and desalted using Nap-5 columns. This resulted in several clusters of Na^+ adducts being observed. Since these experiments required large amounts of material, a thorough optimisation of the experimental conditions was not deemed appropriate. Detailed in the appendices (appendix 9-11) are the raw data and zoomed in regions of the deconvoluted mass spectrum for each peak of interest.

*Table 4.7 - The observed ions from the ESI mass spectrum of **S1TT•S2TT** and **S1FcTT$_{SS}$•S2TT** in the presence of 2 molar equivalents of Hg(ClO$_4$)$_2$.*

	Observed Mass / Da (Predicted Mass)		
	Duplex	**+ 1 T-Hg-T**	**+ 2 T-Hg-T**
S1TT•S2TT	9741 (9743)	9941 (9941)	10139 (10139)
S1FcTTSS•S2TT	9832 (9832)	10030 (10030)	10228 (10228)

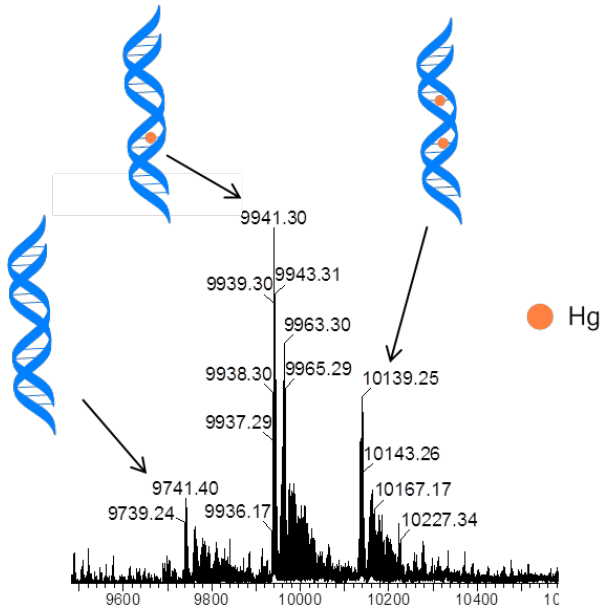


*Figure 4.21 – Deconvuluted ESI mass spectrum observed for **S1TT•S2TT** when bound to mercury showing the three main ions with zero, one and two mercury atoms bound to the duplex.*

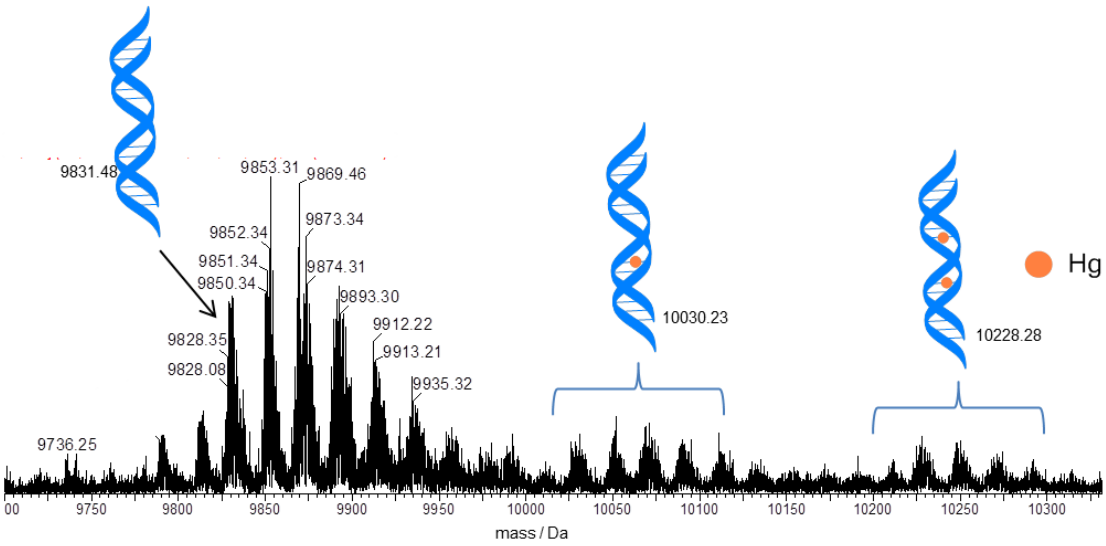


*Figure 4.22 – Deconvuluted ESI mass spectrum observed for **S1Fcn3c2TT$_{SSpRp}$R•S2TT** when bound to mercury showing the three main ions with zero, one and two mercury atoms bound to the duplex.*

These findings suggest that the T-Hg-T interaction is stable enough to withstand the ionisation process that is similar to the results observed by Ono *et al.*_ENREF_12[12] It also supports the UV-vis spectroscopic data that suggests that both T-T mismatches are involved in binding.

4.2.4.5 Electrochemistry

Unfortunately it was not possible to undertake a detailed study of the effects of mercury addition to FcNA-DNA conjugate systems by electrochemistry. The working concentrations for studying this strand (50 μM) in solution to distinguish current from the background (Chapter 3) meant that working concentrations of mercury also had to be high. As a result the addition of mercury(II) perchlorate to solutions at these concentrations resulted in a significant interference of the ferrocene current signal due to mercury adsorption at the electrode surface (Figure 4.23).

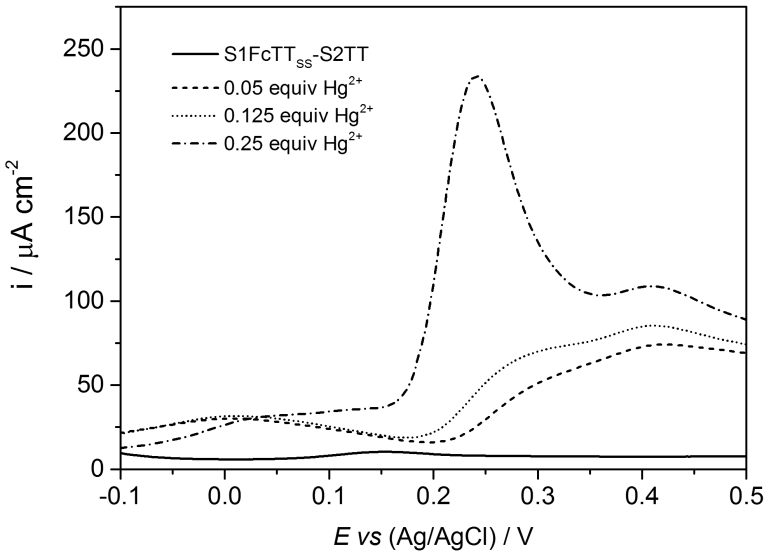


*Figure 4.23 - A SWV showing the effects of adding Hg^{2+} to a solution of **S1FcTT$_{SS}$•S2TT**. 50 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 1 M $NaClO_4$.*

4.3 Conclusions and Future Work

DNA duplexes containing a **Fcn2c3TT**$_{SSpRpR}$ unit as part of a double T-T mismatch site showed the ability to bind Hg(II) in solution as evidenced by a marked increase in duplex stability, it being significantly more stable than when its paired with its fully complementary strand (**S2AA**). This is supported by CD and UV-vis titration studies, with the latter indicating the binding of two Hg(II) ions per duplex. Furthermore mass spectrometry results gave further evidence for the two-site binding of two mercury ions to the T-T mismatches. The system appears to be highly specific for mercury, with other heavy and/or divalent metals showing little changes to CD or UV-vis spectra, indicating weak or no binding at the T-T mismatch sites. Whilst the incorporation of the FcNA motif results in smaller changes when compared with unmodified DNA systems, it does represent what appears to be the first example of an XNA unit being shown to bind mercury. It is anticipated that the redox properties of FcNA labelled oligonucleotides will be able to be fully exploited the future design of electrochemical mercury sensors that could be used to assess a range of complex sample matrices.

However the results here indicate that future work should focus on surface-based sensors. Surface immobilisation would alleviate the adsorption effects that were observed when mercury was added to strands in solution. The closely packed self-assembled monolayers (Au-S) utilised often in DNA electrochemistry would be expected to provide a barrier for the interaction of mercury with the electrode as shown in many literature examples.[78-86] Furthermore surface immobilisation would be important for ensuring that the limit of mercury detection is low and also allow for potential re-use of these sensor devices.

4.4 References

1. H.-A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2003, **42**, 3204-3206.
2. C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. H. Bossmann, N. J. Turro and J. K. Barton, *Science*, 1993, **262**, 1025-1029.
3. G. H. Clever, Y. Söltl, H. Burks, W. Spahl and T. Carell, *Chem. Eur. J*, 2006, **12**, 8708-8718.
4. G. H. Clever, C. Kaul and T. Carell, *Angew. Chem. Int. Ed.*, 2007, **46**, 6226-6236.
5. G. H. Clever and M. Shionoya, *Coord. Chem. Rev.*, 2010, **254**, 2391-2402.
6. S. Katz, K. Gutfreund, S. Shulman and J. D. Ferry, *J. Am. Chem. Soc.*, 1952, **74**, 5706-5709.
7. C. A. Thomas, *J. Am. Chem. Soc.*, 1954, **76**, 6032-6034.
8. S. Katz, *Biochim. Biophys. Acta Biochim Biophys Acta*, 1963, **68**, 240-253.
9. D. W. Gruenwedel and M. K. Cruikshank, *Nucleic Acids Res.*, 1989, **17**, 9075-9086.
10. E. Buncel, C. Boone, H. Joly, R. Kumar and A. R. Norris, *J. Inorg. Biochem.*, 1985, **25**, 61-73.
11. Z. Kuklenyik and L. G. Marzilli, *Inorg. Chem.*, 1996, **35**, 5654-5662.
12. Y. Miyake, H. Togashi, M. Tashiro, H. Yamaguchi, S. Oda, M. Kudo, Y. Tanaka, Y. Kondo, R. Sawa, T. Fujimoto, T. Machinami and A. Ono, *J. Am. Chem. Soc.*, 2006, **128**, 2172-2173.
13. Y. Tanaka, S. Oda, H. Yamaguchi, Y. Kondo, C. Kojima and A. Ono, *J. Am. Chem. Soc.*, 2007, **129**, 244-245.
14. M. Sundaralingam and J. A. Carrabine, *Biochemistry*, 1971, **10**, 292-299.
15. T. Kozasa, Y. Miyakawa, A. Ono and H. Torigoe, *Nucleic Acids Symp. Ser.*, 2008, **52**, 197-198.
16. I. Okamoto, K. Iwamoto, Y. Watanabe, Y. Miyake and A. Ono, *Angew. Chem. Int. Ed.*, 2009, **48**, 1648-1651.
17. T. Dairaku, K. Furuita, H. Sato, J. Sebera, D. Yamanaka, H. Otaki, S. Kikkawa, Y. Kondo, R. Katahira, F. Matthias Bickelhaupt, C. Fonseca Guerra, A. Ono, V. Sychrovsky, C. Kojima and Y. Tanaka, *Chem. Commun.*, 2015, **51**, 8488-8491.
18. H. Yamaguchi, J. ebera, J. Kondo, S. Oda, T. Komuro, T. Kawamura, T. Dairaku, Y. Kondo, I. Okamoto, A. Ono, J. V. Burda, C. Kojima, V. Sychrovsky and Y. Tanaka, *Nucleic Acids Res.*, 2014, **42**, 4094-4099.
19. J. Šebera, J. Burda, M. Straka, A. Ono, C. Kojima, Y. Tanaka and V. Sychrovský, *Chem. Eur. J*, 2013, **19**, 9884-9894.

20. J. Kondo, T. Yamada, C. Hirose, I. Okamoto, Y. Tanaka and A. Ono, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 2385-2388.
21. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
22. A. Ono, S. Cao, H. Togashi, M. Tashiro, T. Fujimoto, T. Machinami, S. Oda, Y. Miyake, I. Okamoto and Y. Tanaka, *Chem. Commun.*, 2008, 4825-4827.
23. A. Ono, H. Torigoe, Y. Tanaka and I. Okamoto, *Chem. Soc. Rev.*, 2011, **40**, 5855-5866.
24. J. S. Lee, L. J. P. Latimer and R. S. Reid, *Biochem. Cell Biol.*, 1993, **71**, 162-168.
25. P. Aich, S. L. Labiuk, L. W. Tari, L. J. T. Delbaere, W. J. Roesler, K. J. Falk, R. P. Steer and J. S. Lee, *J. Mol. Biol.*, 1999, **294**, 477-485.
26. A. Rakitin, P. Aich, C. Papadopoulos, Y. Kobzar, A. S. Vedeneev, J. S. Lee and J. M. Xu, *Phys. Rev. Lett.*, 2001, **86**, 3670-3673.
27. S. D. Wettig, C. Z. Li, Y. T. Long, H. B. Kraatz and J. S. Lee, *Anal. Sci.*, 2003, **19**, 23-26.
28. Y. T. Long, C. Z. Li, T. C. Sutherland, H. B. Kraatz and J. S. Lee, *Anal. Chem.*, 2004, **76**, 4059-4065.
29. X. Li, J. S. Lee and H.-B. Kraatz, *Anal. Chem.*, 2006, **78**, 6096-6101.
30. C. Z. Li, Y. T. Long, H. B. Kraatz and J. S. Lee, *J. Phys. Chem. B*, 2003, **107**, 2291-2296.
31. M. Roitzsch and B. Lippert, *J. Am. Chem. Soc.*, 2004, **126**, 2421-2424.
32. S. S. Alexandre, J. M. Soler, L. Seijo and F. Zamora, *Phys. Rev. B*, 2006, **73**, 2051121-20511215.
33. M. Korbas, S. R. Blechinger, P. H. Krone, I. J. Pickering and G. N. George, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 12108-12112.
34. H. H. Harris, I. J. Pickering and G. N. George, *Science*, 2003, **301**, 1203-1203.
35. M. E. Ariza and M. V. Williams, *J. Biochem. Mol. Toxicol.*, 1999, **13**, 107-112.
36. I. Onyido, A. R. Norris and E. Buncel, *Chem. Rev.*, 2004, **104**, 5911-5930.
37. P. B. Tchounwou, W. K. Ayensu, N. Ninashvili and D. Sutton, *Environ. Toxicol.*, 2003, **18**, 149-175.
38. P. Grandjean, P. Weihe, R. F. White and F. Debes, *Environ. Res.*, 1998, **77**, 165-172.
39. E. M. Nolan and S. J. Lippard, *Chem. Rev.*, 2008, **108**, 3443-3480.
40. Y. Li, C. Chen, B. Li, J. Sun, J. Wang, Y. Gao, Y. Zhao and Z. Chai, *J. Anal. At. Spectrom.*, 2006, **21**, 94-96.

41. F. Han, W. D. Patterson, Y. Xia, B. B. M. Sridhar and Y. Su, *Water, Air, Soil Pollut.*, 2006, **170**, 161-171.
42. K. W. Jackson and G. Chen, *Anal. Chem.*, 1996, **68**, 231-256.
43. N. Kaur, G. Dhaka and J. Singh, *New J. Chem.*, 2015.
44. K. Bera, A. K. Das, M. Nag and S. Basak, *Anal. Chem.*, 2014, **86**, 2740-2746.
45. E. M. Nolan, M. E. Racine and S. J. Lippard, *Inorg. Chem.*, 2006, **45**, 2742-2749.
46. A. B. Descalzo, R. Martínez-Máñez, R. Radeglia, K. Rurack and J. Soto, *J. Am. Chem. Soc.*, 2003, **125**, 3418-3419.
47. X. Guo, X. Qian and L. Jia, *J. Am. Chem. Soc.*, 2004, **126**, 2272-2273.
48. S. H. Mashraqui, S. A. Tripathi, S. S. Ghorpade and S. R. Britto, *Beilstein J. Org. Chem.*, 2012, **8**, 1352-1357.
49. R. Shunmugam, G. J. Gabriel, C. E. Smith, K. A. Aamer and G. N. Tew, *Chem. Eur. J*, 2008, **14**, 3904-3907.
50. M. J. Choi, M. Y. Kim and S.-K. Chang, *Chem. Commun.*, 2001, 1664-1665.
51. E. Palomares, R. Vilar and J. R. Durrant, *Chem. Commun.*, 2004, 362-363.
52. E. Coronado, J. R. Galán-Mascarós, C. Martí-Gastaldo, E. Palomares, J. R. Durrant, R. Vilar, M. Gratzel and M. K. Nazeeruddin, *J. Am. Chem. Soc.*, 2005, **127**, 12351-12356.
53. M. a. Alfonso, A. Tárraga and P. Molina, *J. Org. Chem.*, 2011, **76**, 939-947.
54. F. Zapata, A. Caballero, P. Molina and A. Tárraga, *Sensors*, 2010, **10**, 11311-11321.
55. M. H. Mashhadizadeh and I. Sheikhshoaie, *Talanta*, 2003, **60**, 73-80.
56. D. Huang, T. Hu, N. Chen, W. Zhang and J. Di, *Anal. Chim. Acta*, 2014, **825**, 51-56.
57. G.-W. Wu, S.-B. He, H.-P. Peng, H.-H. Deng, A.-L. Liu, X.-H. Lin, X.-H. Xia and W. Chen, *Anal. Chem.*, 2014, **86**, 10955-10960.
58. Y. Zhao and Z. Zhong, *J. Am. Chem. Soc.*, 2006, **128**, 9988-9989.
59. J.-S. Lee, M. S. Han and C. A. Mirkin, *Angew. Chem. Int. Ed.*, 2007, **46**, 4093-4096.
60. Y. Liu, A. Kuzuya, R. Sha and J. Guillaume, *J. Am. Chem. Soc.*, 2008, **130**, 10882-10883.
61. C.-W. Liu, Y.-T. Hsieh, C.-C. Huang, Z.-H. Lin and H.-T. Chang, *Chem. Commun.*, 2008, 2242-2244.
62. C.-K. Chiang, C.-C. Huang, C.-W. Liu and H.-T. Chang, *Anal. Chem.*, 2008, **80**, 3716-3721.
63. A. Ono and H. Togashi, *Angew. Chem. Int. Ed.*, 2004, **43**, 4300-4302.
64. C.-W. Liu, C.-C. Huang and H.-T. Chang, *Anal. Chem.*, 2009, **81**, 2383-2387.
65. Z. Wang, J. H. Lee and Y. Lu, *Chem. Commun.*, 2008, 6005-6007.

66. Y. Miyake and A. Ono, *Tetrahedron Lett.*, 2005, **46**, 2441-2443.
67. R. Métivier, I. Leray and B. Valeur, *Chem. Eur. J*, 2004, **10**, 4480-4490.
68. N.-N. Bui, J.-T. Hong, S.-i. Mho and H.-Y. Jang, *Bull. Korean Chem. Soc.*, 2008, **29**, 1395-1398.
69. T. Ihara, Y. Maruo, S. Takenaka and M. Takagi, *Nucleic Acids Res.*, 1996, **24**, 4273-4280.
70. H. Cai, X. Cao, Y. Jiang, P. He and Y. Fang, *Anal. Bioanal. Chem.*, 2003, **375**, 287-293.
71. A. Anne, A. Bouchardon and J. Moiroux, *J. Am. Chem. Soc.*, 2003, **125**, 1112-1113.
72. C. E. Immoos, S. J. Lee and M. W. Grinstaff, *ChemBioChem*, 2004, **5**, 1100-1103.
73. M. Inouye, R. Ikeda, M. Takase, T. Tsuri and J. Chiba, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 11606-11610.
74. J. Chiba, A. Akaishi, R. Ikeda and M. Inouye, *Chem. Commun.*, 2010, **46**, 7563-7565.
75. R. Ikeda, S. Kitagawa, J. Chiba and M. Inouye, *Chem. Eur. J*, 2009, **15**, 7048-7051.
76. I. Willner and M. Zayats, *Angew. Chem. Int. Ed.*, 2007, **46**, 6408-6418.
77. B. R. Baker, R. Y. Lai, M. S. Wood, E. H. Doctor, A. J. Heeger and K. W. Plaxco, *J. Am. Chem. Soc.*, 2006, **128**, 3138-3139.
78. R.-G. Cao, B. Zhu, J. Li and D. Xu, *Electrochem. Commun.*, 2009, **11**, 1815-1818.
79. S.-J. Liu, H.-G. Nie, J.-H. Jiang, G.-L. Shen and R.-Q. Yu, *Anal. Chem.*, 2009, **81**, 5724-5730.
80. Z. Zhu, Y. Su, J. Li, D. Li, J. Zhang and S. Song, *Anal. Chem.*, 2009, **81**, 7660-7666.
81. D. Han, Y.-R. Kim, J.-W. Oh, T. H. Kim, R. K. Mahajan, J. S. Kim and H. Kim, *The Analyst*, 2009, **134**, 1857-1862.
82. D. Wu, Q. Zhang, X. Chu, H. Wang, G. Shen and R. Yu, *Biosensors Bioelectron.*, 2010, **25**, 1025-1031.
83. J. Zhuang, L. Fu, D. Tang, M. Xu, G. Chen and H. Yang, *Biosensors Bioelectron.*, 2013, **39**, 315-319.
84. G. V. Guerreiro, A. J. Zaitouna and R. Y. Lai, *Anal. Chim. Acta*, 2014, **810**, 79-85.
85. P. Miao, L. Liu, Y. Li and G. Li, *Electrochem. Commun.*, 2009, **11**, 1904-1907.
86. Z. Zhang, A. Tang, S. Liao, P. Chen, Z. Wu and G. Shen, *Biosens. Bioelectron.*, 2011, **26**, 3320-3324.
87. http://www.epa.gov/mercury/regs.htm, Accessed 01/06/2015, 2015.
88. D. W. Gruenwedel, *J. Inorg. Biochem.*, 1994, **56**, 201-212.
89. J. Eisinger, R. G. Shulman and B. M. Szymanski, *J. Chem. Phys.*, 1962, **36**, 1721-1729.

90. E. E. Kriss and K. B. Yatsimirskii, *Russ. Chem. Rev.*, 1966, **35**, 144-152.
91. P. W. Schneider, H. Brintzinger and H. Erlenmeyer, *Helv. Chim. Acta*, 1964, **47**, 992-1002.
92. E. Frieden and J. Alles, *J. Biol. Chem.*, 1958, **230**, 797-804.
93. M. Zhang and B.-C. Ye, *Analyst*, 2011, **136**, 5139-5142.

Chapter 5

Redox-Active Metal Macrocycles in DNA

The work detailed in this chapter was completed in collaboration between the Tucker Group and Dr. Jarosław Kowalski (Polish Academy of Sciences). Parts of the results shown within were completed with or by Dr. Jarosław Kowalski or Dr. Jean-Louis Duprey. J.K was responsible for the synthesis of the macrocyclic monomers for incorporation into DNA and J-L.D undertook much of the thermal melting analysis and modelling studies.

5.1 Introduction

Oligonucleotides that incorporate signal-transducing moieties capable of providing information on minute changes in structure in target DNA strands have been of much interest in recent years. The drive towards personalised medicines and identification of bio-markers for disease has seen a number of DNA modifications designed to allow changes in genetic sequence to be monitored and epigenetic changes, such as the prominence of methylated nucleobases to be sensed. Additionally, interest in molecular computing and nano-electronics has led to vast amounts of research into modified DNA as potential nano-wires,[1] where various changes in structure have been made for probing the conducting behaviour of DNA duplexes.

5.1.1 Single Nucleotide Polymorphisms

The development of genomic sequencing has rapidly expanded our understanding of the human genome; techniques are now at the stage that a single genome can be sequenced in a matter of hours. The exploration of this information within the genome has raised the significance of the identity of small genetic changes in the DNA code between individuals. Single Nucleotide Polymorphisms (SNPs), the most common form of genetic change, are variations in DNA sequences where a nucleobase at specific locus differs (see Figure 5.1) among >1% of the population. SNPs are extremely common within mammalian DNA, where the majority are found in non-coding rather than coding regions. Of those that occur within the coding region, there are two types of SNP: synonymous, where a SNP directly affects the protein sequence within a protein, and nonsynonymous, where the SNP results in no change in protein synthesis. The occurrence of a protein sequence change as a result of a SNP is a result of degeneracy within DNA and the 3-base codon recognition machinery of ribosomes.

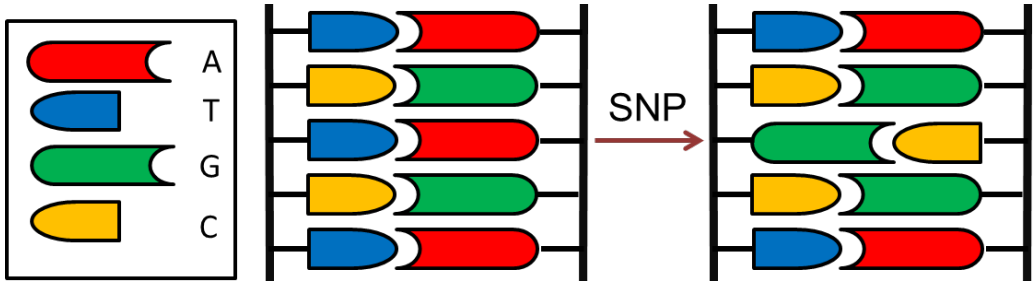


Figure 5.1 - An example of a SNP where an A-T base pair becomes G-C.

In recent years certain SNPs have been found to be associated with diseases such as Alzheimer's[2] and cancers,[3] and efforts are now being made to identify these SNPs in individuals at an earlier stage. The design of probes for the detection of these variations has resulted in many new technologies being developed including now commercially available assays such as Taqman[4] and Scorpion probes.[5] Unfortunately current methods available for SNP genotyping are considered to be somewhat cumbersome and usually require a high degree of technical expertise. Therefore, personalised medicine becoming more prevalent in an ever-expanding global population, the need for rapid, easy to use and inexpensive assays for evaluating human health has never been greater. They tend to rely on binding differences between fully matched and mismatching sequences. This means that the probes can only operate in a narrow and elevated temperature window with the conditions highly specific for each sequence of interest.

5.1.2 DNA Intercalators

The interaction of molecules with nucleic acids has for some time been of interest to the supramolecular chemist. Recognition of specific genetic elements of the DNA/RNA structure holds promise for switching on or off gene expression. Interactions with nucleic acids can occur via five main pathways: major groove binding; minor groove binding; covalent binding or coordination to the bases; recognition of the negatively charged sugar-phosphate

backbone and intercalation between the nucleobase stack.[6] The focus of this discussion is with the use of intercalators for binding.

Intercalators by definition are capable of inserting themselves into the hydrophobic core of the DNA/RNA duplex and stack between nucleobase pairs through π-π interactions. Intercalators are planar, aromatic motifs, and, often bear a positive charge, which aids DNA binding as a result of intercalation bringing the group into close proximity to the anionic backbone. Intercalators have been traditionally studied for ligand-DNA interactions as potential therapeutics when free in solution (Figure 5.2a). More recently, a particular class of DNA modification, involving the tagging of strands with intercalating groups, has shown to be effective for providing information on DNA structural changes, whether through spectroscopic or electrochemical means. Their incorporation into DNA *via* automated DNA synthesis has become more common, giving better control over stoichiometries of the intercalation process (Figure 5.2). Intercalating compounds are generally good chromophores owing to their aromaticity; however, the ability to transduce signals that are distinguishable from the background makes certain designs more desirable than others from both a research and an application perspective. Two signal transducing properties are described below.

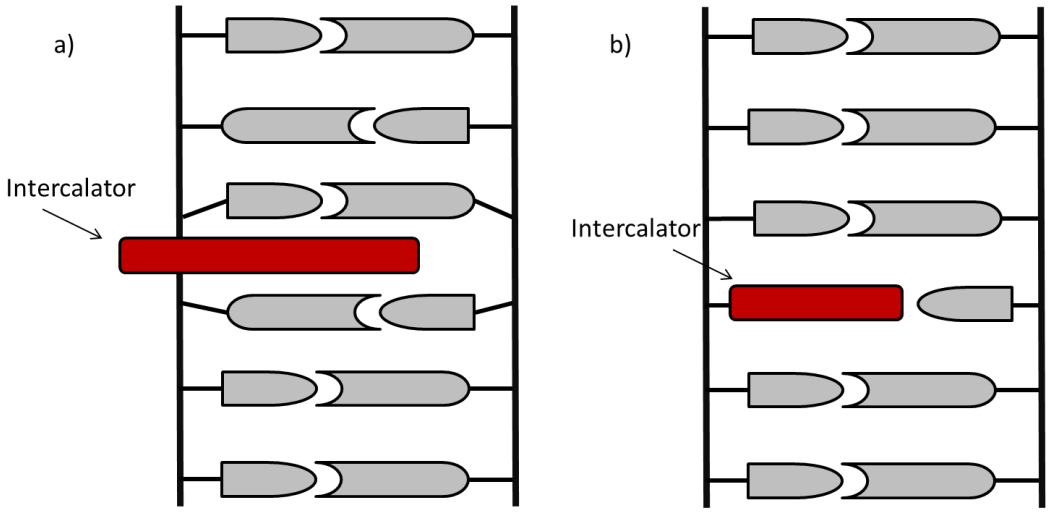


Figure 5.2 - A schematic representation of the two intercalation approaches. a) shows the intercalation of free species in solution while b) shows the incorporation of an intercalator into the DNA backbone via DNA synthesis.

5.1.2.1 Fluorescent Intercalators

The use of inherently fluorescent organic molecules such as anthracene,[7,8] pyrene[9] and thiazole orange[10,11] has shown great promise for the development of SNP sensing hybridisation assays, and are termed base discriminating fluorophores (BDF's). By incorporation of such moieties into the backbone of DNA or artificial nucleic acids such as PNA, such probes have been shown to discriminate single base changes in target DNA sequences. Whether through changes in fluorescence emission intensity or via Stokes shifts, these intercalating motifs are designed so that they can slot directly between stacking base pairs to probe directly the specific changes in local environments.

In one of the first examples described by Saito *et al*, pyrene-modified DNA could be used to identify the base opposite its intercalation pocket as a result of the fluorescence emission changes. The modification was incorporated *via* a Suzuki-like coupling to the thymine nucleobase before incorporation into DNA *via* automated DNA synthesis. In addition, the group also described anthracene-based intercalators that display similar base discriminating behaviour. Some examples of these modifications are shown in Figure 5.3**.**

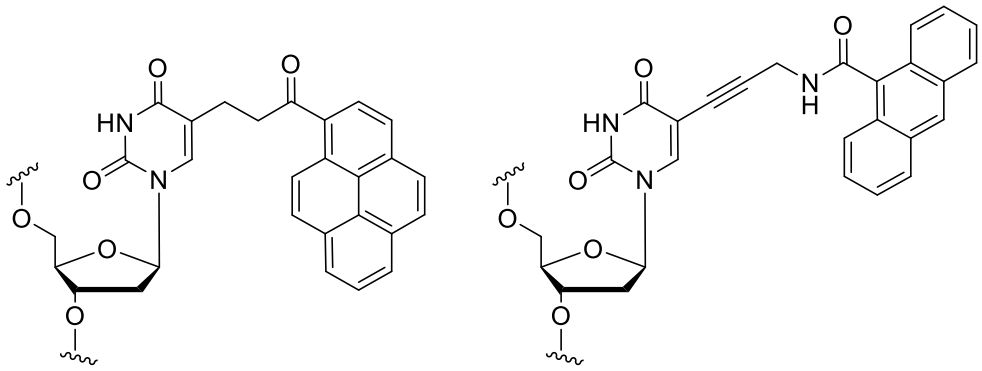


Figure 5.3 - Examples of organic base discriminating fluorophores developed by Saito. Left – pyrene.[9] Right – anthracene.[7]

Similarly, the Tucker group designed anthracene-based DNA probes with the ability to detect SNPs in a range of sequences, including those associated with prostate cancer and

Alzheimer's disease.[8,12,13] In contrast to Saito's modifications, these anthracene-based probes involve the deletion of the nucleotide unit from the probe sequence and incorporation of the intercalating unit *via* a threoninol linkage (shown in Figure 5.4), in which the stereochemistry and hence sensing properties can be fine-tuned.

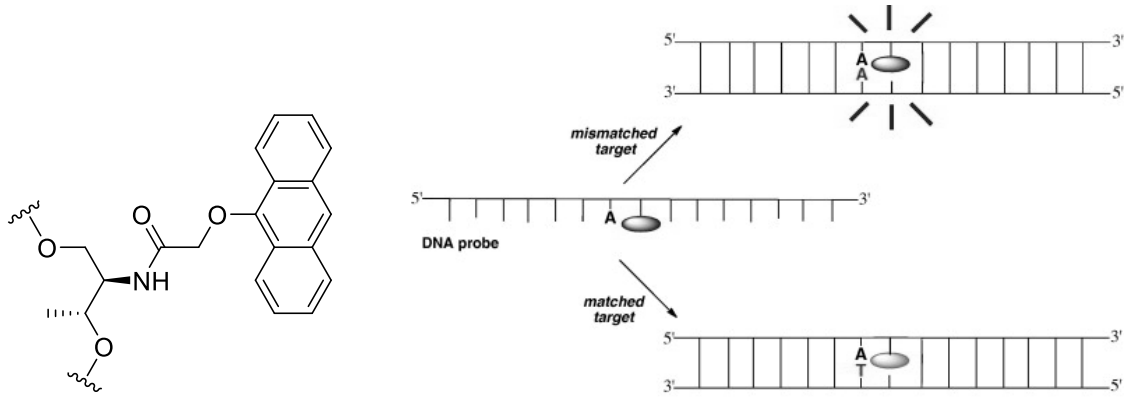


Figure 5.4 - Tucker groups anthracene SNP detecting fluorophore which has the ability to discriminate between SNP sequences via changes in fluorescence emission.[8,12]

While these fluorescence approaches have proved successful in proof-of-concept research, there are number of challenges that are associated with their design. The fluorescence emission intensity of a fluorophore is often sensitive to changes in salt concentration and temperature. Additionally, the use of fluorophores that emit at higher energies, such as anthracene and pyrene, can be hindered by the observed background fluorescence in biological samples. While this is not an issue in post-PCR amplication assays, it does pose a number of challenges for directly probing living cells or for potential use in on-site diagnostic devices.

5.1.2.2 Redox-Active Intercalators

Another inherent property of some intercalators that can be exploited is their ability to undergo reversible electron transfer. While many redox-active DNA probes that have been described to date are non-intercalating ferrocene labelled motifs (Chapter 3), there has been progress in the development of planar aromatic motifs such as methylene blue (MB) and

anthraquinones,[14,15] whose redox properties are well understood. In particular, methylene blue, owing to its electrochemical stability, optical properties and well understood intercalation properties[16] has become a focus of many research groups.[17-24] Earlier work focussed on MB's ability to intercalate into DNA by addition in solution to a surface immobilised duplex. Barton *et al.* reported on an electrochemical assay (shown in Figure 5.5) which could discriminate between canonical nucleobases A, G, C and T,[21] as well as distinguish between base analogues and DNA damage products.[25] The mechanism for this discriminative behaviour has been related to changes in the mediated charge transport (CT) network afforded by the π-stack of the DNA duplex, which can be monitored directly by the intercalation of MB.[21,25] These initial studies were, however, limited in that intercalation was susceptible to non-specific binding and the requirement for additional exogenous reagent control as well as washing steps.

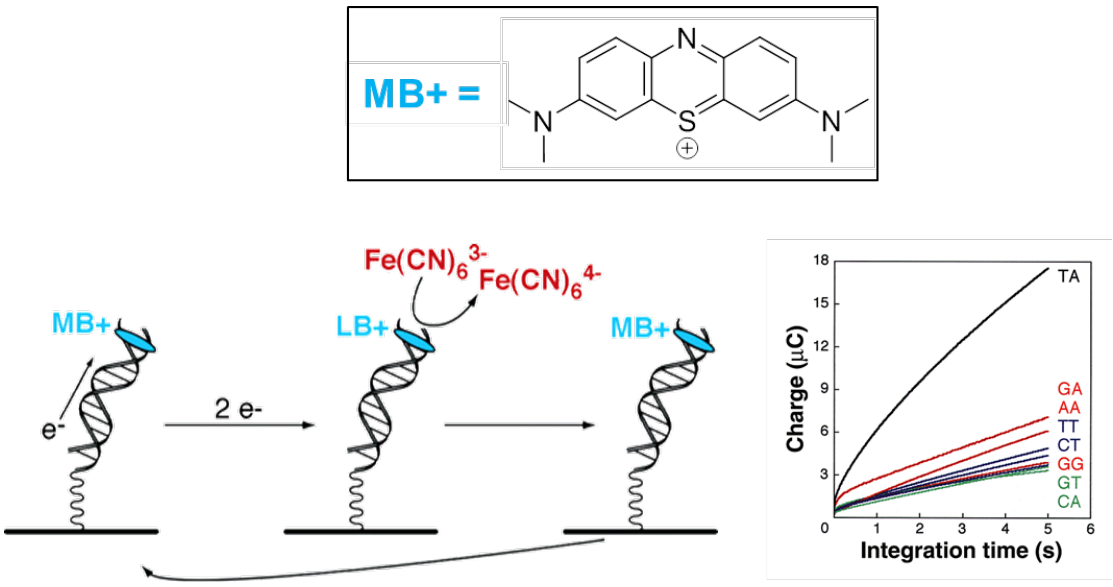


Figure 5.5 - MB SNP sensing redox probes published by Barton et al.[21,25] Charge transfer changes afforded by SNP sites are transduced by the electrochemical output of the intercalated MB. Edited from ref.[21]

More recently, the use of methylene blue has focussed on its application as an electrochemical tag for monitoring large bulk changes in DNA structure, such as the

formation of duplexes and stem-loops. The direct incorporation of MB into oligonucleotides *via* post-synthetic labelling and its enhanced electrochemical stability over ferrocene-based tags (see Chapter 3) has made it a significantly more popular choice in this respect. The groups of Barton and Plaxco have developed assays (see Figure 5.6) capable of discriminating between different oligonucleotide sequences and even to the sensing of biologically relevant molecules using DNA aptamers[26-28] (Figure 5.7). While SNP-based sensors have been developed using labelled oligonucleotides, their sensing behaviour remains largely based on large bulk conformational changes rather than transducing information regarding an exact binding site, *i.e.* a mismatch site. Furthermore, despite post-synthetic labelling of oligonucleotides being possible, it remains expensive and the synthesis of MB and its derivatives is challenging. Only very recently (early Feb 2015) has a MB derivative capable of being used in automated phosphoramidite synthesis been developed. Even so, modifications to the structure of MB and to the oligonucleotide synthesis procedures are still required to prevent its degradation.[29,30] The development of other intercalating redox probes has been hindered by similar synthetic challenges.

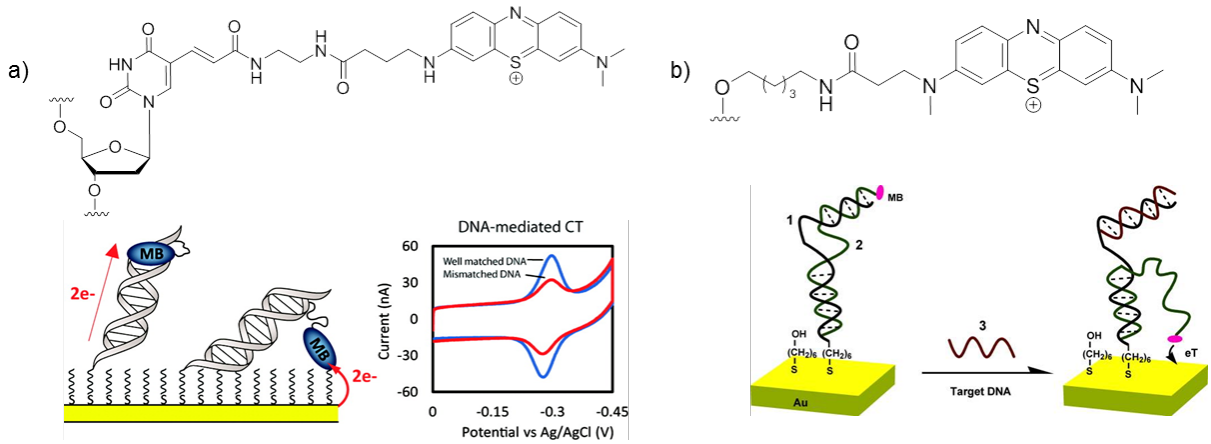


Figure 5.6 - Schematic representations of methylene blue labelled oligonucleotides for electrochemical detection of a) SNPs as published by a) Barton[18] and b) Plaxco.[27] Also shown are the methylene blue derivatives associated with these designs.

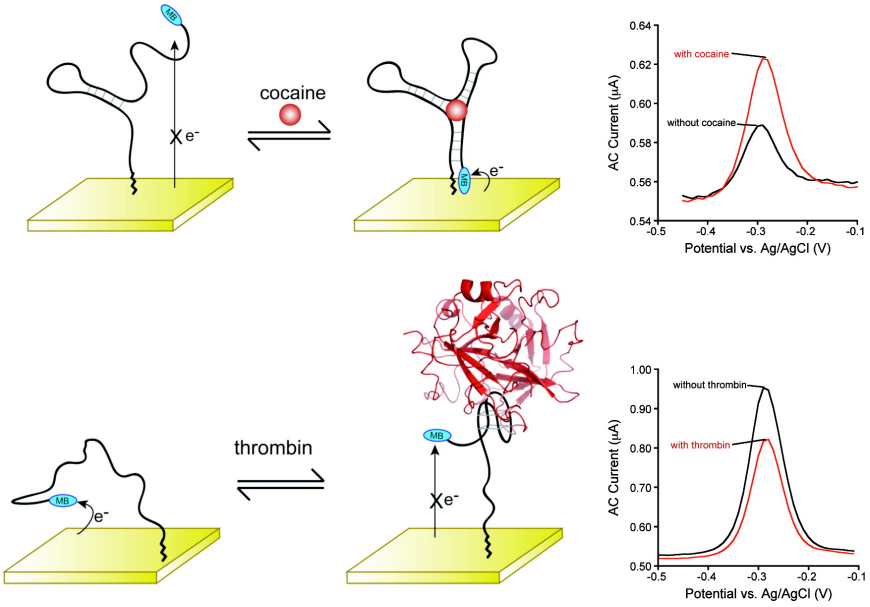


Figure 5.7 - Aptamer based sensing platforms developed by Plaxco and co-workers. Top - Electrochemical aptasensor for the detection of cocaine using methylene blue tagged DNA.[26] Bottom - Signal OFF aptasensor for the detection of thrombin using methylene blue tagged DNA strands.[28] Taken from ref. [31].

Whilst to date there are not many examples of redox-active intercalators that can directly probe subtle changes in a DNA sequence, their use in electrochemical assays remains a promising avenue for the development of rapid, reliable and cheap assays. Electrochemical methods require simple apparatus that can be readily miniaturized and the incorporation of DNA probes onto a surface removes the majority of background interferences and noise associated with techniques based on optical spectroscopy. Furthermore, the potential for probing *in vivo via* electrochemical methods is aided by the rarity of electro-active contaminants found in *in vivo* samples. Plaxco and co-authors share an excellent discussion concerning the use of electrochemical methods for detection in two reviews.[32,33]

5.1.3 Metallo-Intercalators

Whilst the above mentioned intercalators are based upon planar aromatic systems, there remains a significant drive for the design of metallo-complexes as intercalators. Metal centres have the ability to withdraw electron density from their coordinated ligand framework

which increases the electron deficiency of the aromatic systems and in turn enhances π stacking interactions with the nucleobases.[34] Furthermore metal centres can bind to DNA through direct coordination with the nucleobases or negatively charged phosphate backbone. The positive charges that can be imparted onto intercalators using a metal centre also provides the possibility for enhanced binding strengths with DNA.

The first example of such a design described by Lippard *et al* was platinum(II) terpyridine (Figure 5.8a) which was shown to intercalate between the base-stack of DNA base pairs and displayed interesting therapeutic properties.[35,36] The terpyridine moiety and coordinated ligands have been manipulated to control both the binding strength to DNA duplexes, but also to introduce new binding modes to structures such as G-quadruplexes.[37-39] As well as introducing charge, metallo-complexes can be exploited to introduce other interesting properties such as luminescence or electrochemical properties. Octahedral ruthenium and rhodium (such as that shown in Figure 5.8b) complexes have been shown to display fluorescence and DNA cleavage properties respectively.[40] Recently, and relevant to the work described in this chapter, Alison Roger *et al* described the intercalation properties of nickel(II) and copper(II) tetraza-macrocyclic complexes.[41]

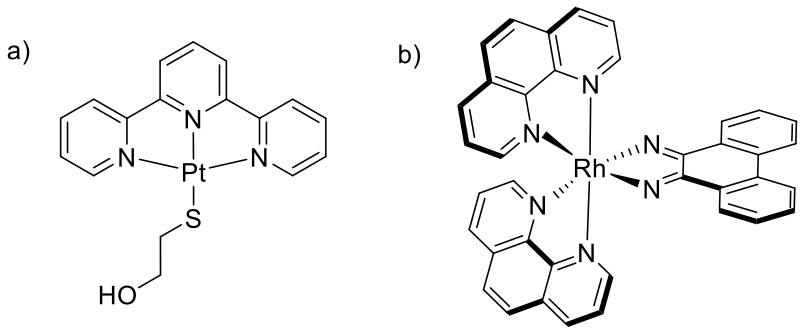


Figure 5.8 – Examples of Metallo-intercalators a) $[Pt(tpy)(SCH_2CH_2OH)]^+$ and b) $[Rh(phen)_2phi]^{3+}$ described by Lippard[35] and Barton respectively.[40]

5.1.4 Metal Complexes Incorporated into DNA

The incorporation of metal complexes into DNA is a highly attractive field of study, given the vast range of potential applications such as nano-wires for computing and electronics, redox-labelling and electrochemical sensing. Incorporation of metals into DNA can be carried out using three main approaches:

1. Post DNA synthesis labelling with metal complexes.[42]
2. Incorporation of a ligand or metal chelator through DNA synthesis and subsequent addition of a metal.[43]
3. Direct incorporation of a complex *via* automated DNA synthesis or enzymatic incorporation.

Unfortunately, due to instability under DNA synthesis conditions, the direct incorporation of metal complexes has been limited to ferrocene compounds (a review on which can be found in ref. [44]) octahededral bipyridyl osmium and ruthenium complexes,[45-51] and lanthanide texaphyrins[52]. More recently Stulz *et al*. described the incorporation of a zinc porphyrin, in which the metal-free macrocycle is incorporated *via* DNA synthesis and the metal is subsequently chelated, bringing about interesting electrochemical properties that can sense duplex formation.[53] Some examples of these incorporations are illustrated in Figure 5.9.

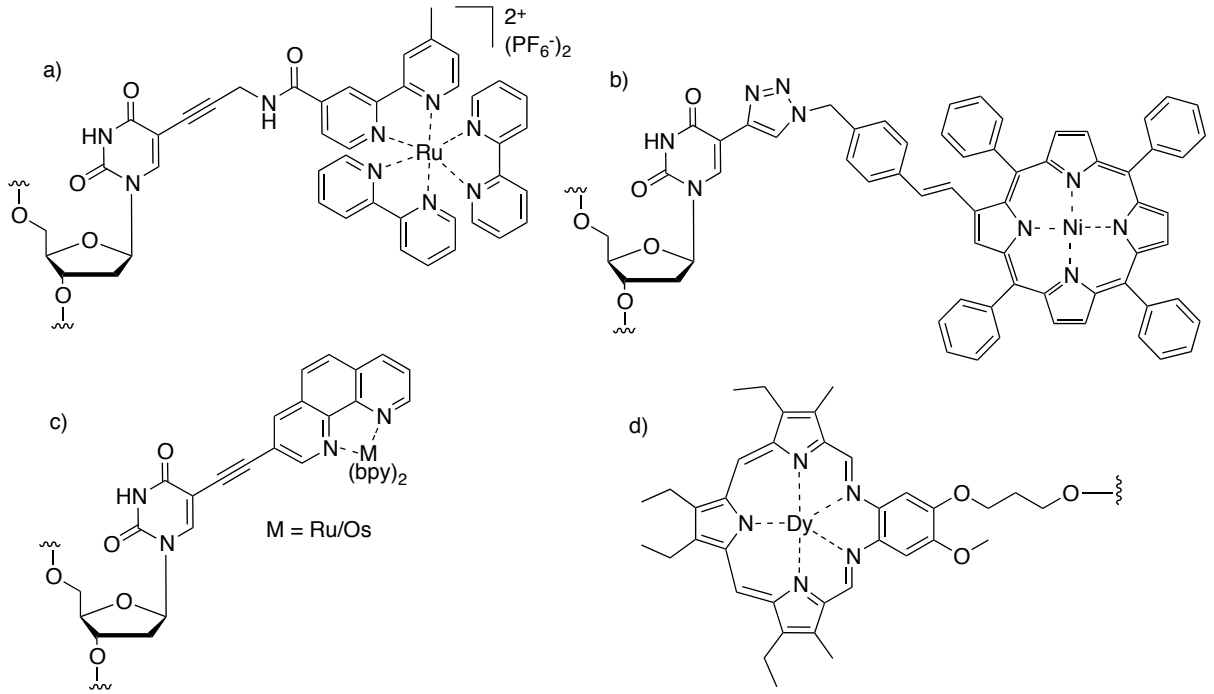


Figure 5.9 - Examples of metal incorporations into DNA. a) trisbipy ruthenium complex [48] b) nickel porphyrin[54] c) octahedral ruthenium and osmium complex[46] and d) dysprosium texaphyrin labels.[55]

Although there are now quite a few examples of DNA tagged with metal complexes there still remains challenges with the use of synthesised monomers capable of withstanding the harsh conditions of DNA synthesis, in particular those metal complexes with geometries and sizes suitable for intercalation into DNA. Without the ability to intercalate, the majority of the complexes publishes have shown an inability to discriminate between subtle changes in the DNA local environment (i.e. for SNP sensing), except for some initial studies on octahedral ruthenium complexes.[45,48,50]

5.1.5 Conclusions and Project Aims

The challenges associated with synthesising modified DNA containing tagged planar metal complexes with the ability to intercalate into DNA and probe the local environment of a duplex are considerable. However, the potential advantage of metal complexes, in terms of their spectroscopic and electrochemical properties makes them ideal candidates for incorporation into DNA as molecular probes. Therefore the aim of this work was to

synthesise a series of stable metal-based macrocyclic complexes that could not only withstand the rigours of automated DNA synthesis and purification but also be able to participate the electrochemical probing of duplex DNA *via* intercalation. By combining the redox properties of tags such as methylene blue and ferrocene with the intercalation properties of pyrene and anthracene derivatives, it was hoped that the metal centre could be used to probe directly the specific environment of the DNA core. Furthermore, by coupling this approach alongside techniques such as CD and UV-vis spectroscopy a new methodology for SNP sensing could be developed and rationalised to compliment previous fluorescent-based approaches by the Tucker Group.[13,56,57]

5.2 Results and Discussion

5.2.1 Design Strategy

Metal centred macrocycles based around tetraimines were proposed as a potential candidate for incorporation into DNA. This class of complexes has been well investigated and their attractive properties such as routine routes to their synthesis, chemical stability and electrochemical activity provided a strong rationale for their incorporation into DNA.[58]

Neutral copper(II) and nickel(II) cyclidene complexes (Figure 5.10) that possess the above mentioned properties[59] as well as strong pi-donor[60] and nanostructural properties[58] were highlighted as possible candidates. These complexes have been shown to be capable of accessing unusual oxidation states (such as Cu(III) and Ni(III)) for transition metals *via* electrochemical oxidation and reduction.[59] Wozniak *et al* have published an extensive review of the properties and applications of these complexes.[58]

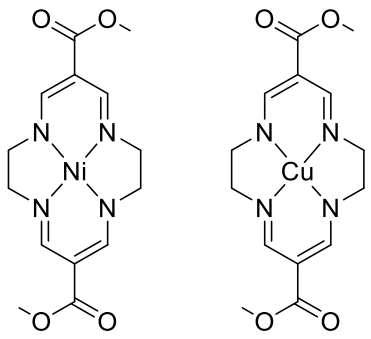


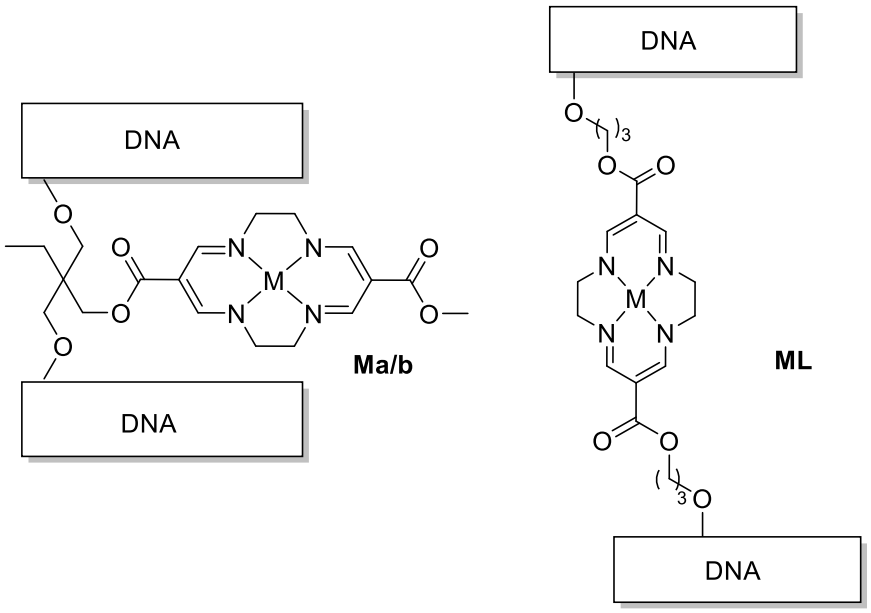
Figure 5.10 - Neutral Ni(II) and Cu(II) tetraazacyclene macrocyles previously reported.[46]

The macrocyclic rings in the complexes exhibit two separate delocalised 7π electron systems, each bearing a -1 charge, which are separated by virtue of the two C_2H_4 aliphatic linkages. Both nickel and copper complexes have been shown by x-ray crystallography to exist in an almost perfect square planar arrangement, with very small deviations (within error) of the metal centre from the ideal plane of the nitrogen donors.[59,61] Additionally both

complexes readily undergo a 1 electron oxidation to the M(III) ion, as shown by cyclic voltammetry. Further discussion of these properties is included later in this discussion. Given the planarity, redox properties and relative ease of synthesis of these complexes they were chosen for incorporation into DNA as potential metal-based SNP sensors.

Our design principles focused on two main categories of compounds (Scheme 5.1):

a) an intercalating system, where a non-nucleosidic base mimic can potentially insert itself into the DNA duplex (M*a/b*).
b) A bridging/linker system as a spacer group that can bridge sections along its DNA backbone. (ML)



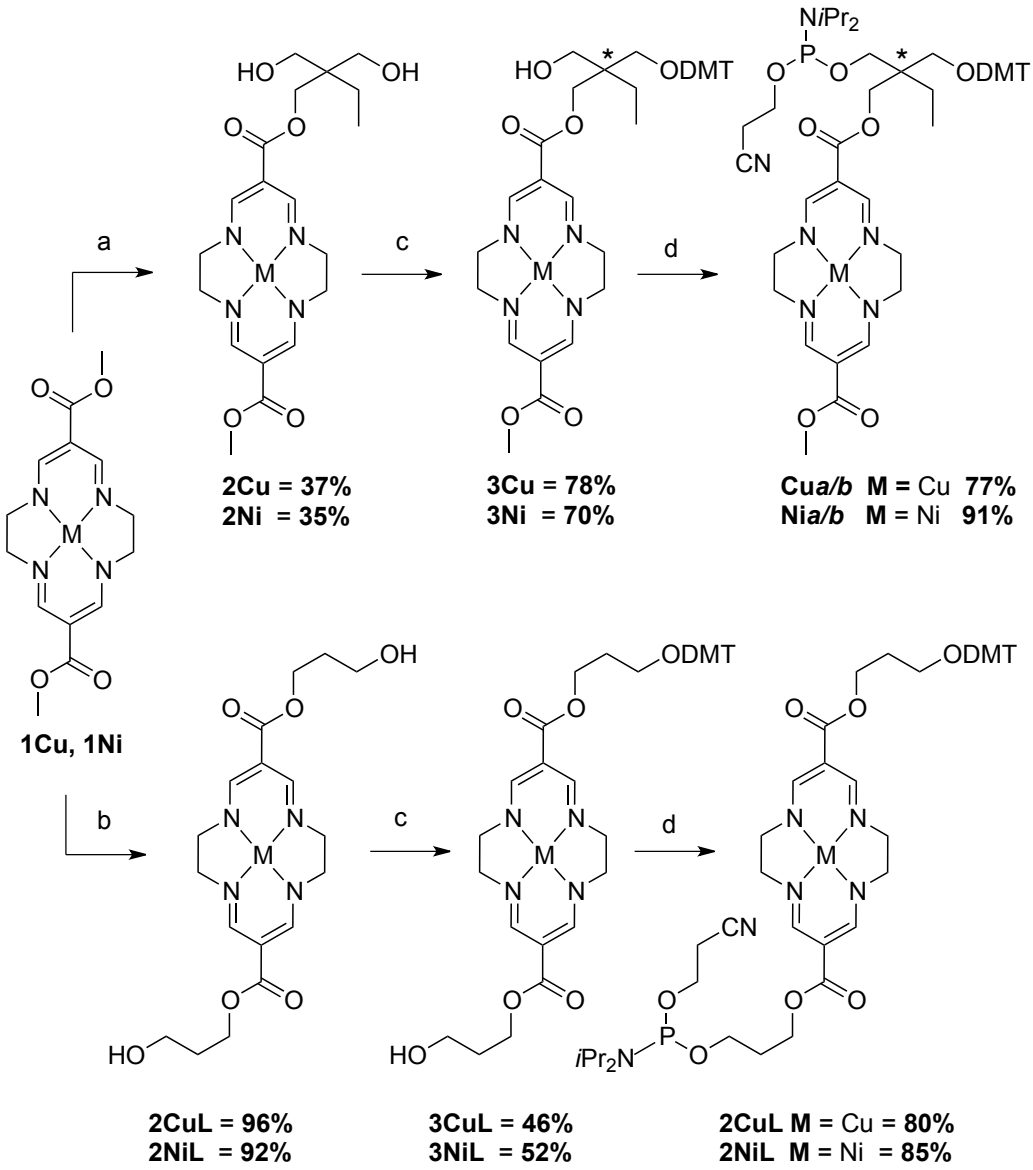
Scheme 5.1 - Diagrammatic representation of the two designs for metal cyclidene macrocycles. Where M = Cu/Ni

5.2.2 Monomer Synthesis

Synthesis of these compounds was carried out by Dr.Jarosław Kowalski at the Polish Academy of Sciences, Warsaw and the University of Birmingham, UK. Full experimental details and characterisation data (J.K) are beyond the scope of this discussion but for interest are included in Experimental section 7.1.2.

In brief, functionalising neutral cyclidenes (Scheme 5.2) was carried out by transesterification of the methyl esters (Linkers) in good yield (>92%). Similarly, synthesising the intercalating moieties (**Cua/b, Nia/b**) is achieved by a base-catalysed transesterification of **1Cu** or **1Ni** with trimethylolpropane. The mono products (**2Cu** and **2Ni**) were then isolated by reverse transesterification in the presence of methanol, as a result of this step the yields are comparably lower (35-37%). The four diols (**3CuL, 3NiL, 3Cu, 3Ni**) were then tritylated and subsequently phosphitylated using previously described chemistries.[62] Note that these steps induce chirality in the intercalating motifs **(3Cu, 3Ni**) which are inseparable prior to conjugation into DNA. The two diastereomers produced are labelled as **Cu*a/b*** and **Ni*a/b*** for the remainder of this chapter. The same synthetic approach was applied to a 'metal-free' control version of **3Cu/3Ni**, in which the metal centre was detemplated prior to the tritylation step using strong acid. Interestingly, and an early indication of the potential stability of these complexes, it was found that the metal-templated macrocycles were readily functionalised with their respective linkages and protecting groups under elevated temperatures (in excess of >80 °C) under basic conditions. Both of which are requirements for success in automated DNA synthesis and purification.

The complexes obtained from x-ray crystallography data showed small deviations (within error) from perfect square planar geometries, suggesting these *quas*i-aromatic systems are indeed suitable for stacking with the base-pairs of a DNA duplex.



Scheme 5.2 - Synthetic route towards metal (Cu/Ni) / no metal centred-macrocycles for incorporation into DNA. (a) 1) trimethylolpropane, Na, 90 °C, 2) MeOH, Na, 80 °C, (b) 1,3-propanediol, Na, 90 °C, (c) DMTrCl, pyridine, RT, (d) chlorodiisopropyl cyanoethylphosphoramidite, DIPEA, DCM,

5.2.3 Incorporation into DNA

5.2.3.1 DNA Synthesis

Incorporation of the modified phosphoramidites was achieved *via* automated DNA synthesis using methodologies previously described by the group. Interestingly, the relative yields of the synthesis and subsequent deprotection steps (30% ammonia solution) appeared to be reliant on both the motif itself and its metal centre. The coupling/trityl yields give an

approximation of the success of the coupling but are not directly related to the stability of the phosphoramidite to the CPG bound oligonucleotide (for more information on DNA synthesis, please refer to Chapter 2.1). Comparisons of the trityl yields of the synthesised oligonucleotides **S1Cu**, **S1Ni** and **S1NoMet** (where no metal ion sits in the macrocycle centre) showed stark differences, which indicate possible degradation of the complex or inherent storage stability differences between the complexes. Incorporation of the copper moiety resulted in an average 5% (± 2%) decrease of the step-wise trityl yield showing remarkably efficient coupling, as opposed to the nickel complex (**3Cu**) which resulted in a 21 % (± 7%) reduction in overall coupling yield. In comparison the metal free complex (**3NoM**) gave a coupling yield in between these two, with a reduction in overall yield of 15 % (± 4%) for this step.

In fact these coupling yields did not reflect the ultimate yields of the oligonucleotide directly the chemical stability of these complexes. The stability of the complexes to post synthesis deprotection using a 30 % ammonia solution (65 °C, 6 hours) varied. This was evidenced by the metal free (**3NoM**) complex producing highly impure and very low yielding products following purification. The reason for this instability is unknown but future studies could investigate alternative methodologies for its synthesis as the ability to add a metal to the macrocyclic cavity could be advantageous for nanoscale technologies. Given the large amounts of material needed for full characterization, and the requirement to investigate these strands using various techniques such as electrochemistry (*circa.* 10 μM solutions), it was the two metal-based complexes that were decided to carry on investigations with.

The sequences containing the modification position are detailed in Table 5.1 below. In order to broaden our understanding of the effect of these modifications, single and multiple incorporations were made into the central region of the strands **S1** and **S2**. These sequences have been well studied within the group[8,13] and have already been discussed in Chapters 3 and 4:

S1X – TGGACTC-**X**-CTCAATG

S2Y – CATTGAG-**Y**-GAGTCCA

Where X = T, TT or macrocyclic incorporation, and Y = A, G, C, T, Abasic or macrocyclic incorporation.

Table 5.1 - Synthesised oligonucleotides bearing macrocylic modifications.

Code	Sequence (5'-3')
S1X[a]	TGGACTC-**X**-CTCAATG
S1NiL	TGGACTC-**NiL**-CTCAATG
S1CuL	TGGACTC-**CuL**-CTCAATG
S1NiLNiL	TGGACTC-**NiLNiL**-CTCAATG
S1CuLCuL	TGGACTC-**CuLCuL**-CTCAATG
S1Ni*a/b*	TGGACTC-**Ni*a/b***-CTCAATG
S1Cu*a/b*	TGGACTC-**Cu*a/b***-CTCAATG
S1NoMet*a/b*[b]	TGGACTCT-**NoMet**-CTCAATG
S1Ni*a/b*Ni*a/b*[c]	TGGACTC-**Ni*a/b*-Ni*a/b*** -CTCAATG
S1Cu*a/b*Cu*a/b*[c]	TGGACTC-**Cu*a/b*-Cu*a/b*** -CTCAATG
S2Y[d]	CATTGAG-**Y**-GAGTCCA
S2NiLNiL	CATTGAG-**NiL**-**NiL**-GAGTCCA
S2CuLCuL	CATTGAG-**CuL**-**CuL**-GAGTCCA

[a] X = T or TT [b] NoMet complex did not survive DNA synthesis.[c] Where multiple incorporations are included the same stereoisomer is used for each incorporation.[d] S2Y, where Y = A, G, C, T, AA or Abasic. Full experimental details and characterisation is detailed in the appendices.

5.2.3.2 Purification

Following the harsh synthesis and deprotection conditions, the modified oligonucleotides were purified using RP-HPLC. Modified oligonucleotides were distinguishable from shorter truncated sequences owing to the increased lipophilicity imparted by the *quasi*-aromatic macrocycles, evident from the shifts in retention time compared with unmodified oligonucleotides. Interestingly, the chirality carried by insertion complexes **Cu*a/b*** and **Ni*a/b*** produced two main peaks of interest in the HPLC traces (an example is shown in Figure 5.11), which allowed for the purification of the two possible diastereomers. These diastereomers had proved difficult to separate following the final synthetic steps and therefore the isolation of each isomer was extremely beneficial to our studies. Being able to separate diastereomers before testing the sensing ability has proved vitally important for the

development of similar intercalation based SNP sensors.[12] Separation of a few minutes allowed for the isolation of isomerically pure strands but a slight disadvantage of this approach, rather than a pre-DNA synthesis separation, is that the overall yield of each isomer is halved. Therefore, multiple syntheses were required to obtain the amounts of each isomer usually obtained for DNA synthesis (1.5 ml approx. 200-300 μM). The difference in retention time of these isomers was not dissimilar to the trends previously observed with serinol and threoninol linked intercalators, providing some insight into which isomer was obtained.[5,9,10] For purposes of this discussion the fastest isomer eluted off the column is defined as **Ni/Cu*a*** and the slower, **Ni/Cu*b***. The linker modified sequences, **Ni/CuL**, bearing no chirality, exhibit only one isomer (an example is shown in appendix 16, page 272).

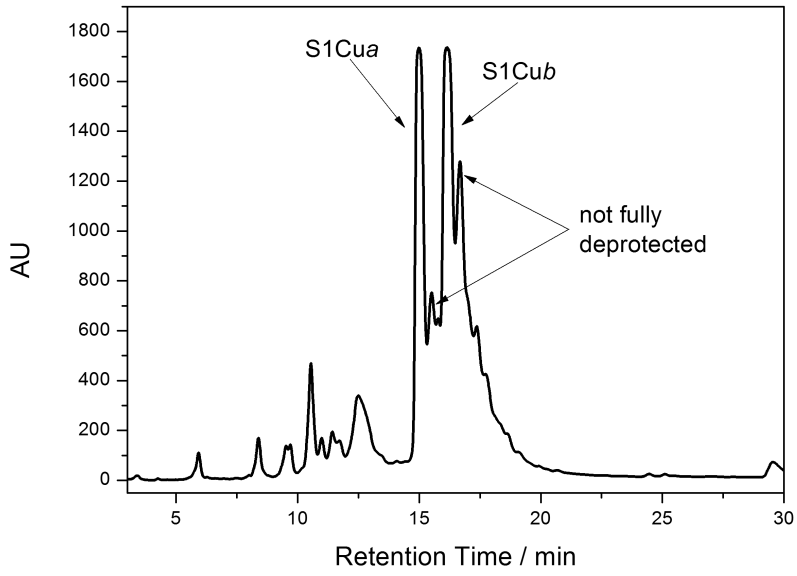


*Figure 5.11 - HPLC trace of copper intercalator (**S1Cua/b**) showing the separation of slow and faster isomers.*

Full characterisation of these strands can be found in appendix 14 (pg. 258). The following sections detail the spectroscopic investigations of the linker systems and the intercalating systems separately.

5.2.4 Linker System

5.2.4.1 UV-Vis Characterisation

Macrocyclic compounds are not often associated with being brightly coloured but the π electron-rich nature of these porphyrin-like analogues leads to some strong transitions in the near-UV region of the spectrum. Similar tetraazatetraene macrocycles, which forms the foundation of these modifications, have been previously shown to absorb strongly in this region. The spectra exhibit two key absorptions, a strong band (300-390 nm) and a shoulder band (330-370 nm), which arise from a Soret band analogous to a π-π* transition and Q bands, respectively.[52] Similarly the modified DNA shows strong absorptions in the same regions (see Figure 5.12). Subtraction of the DNA contributions from the spectrum shows the π-π* transitions in slightly lower energy than previously reported,[61] perhaps as a result of a lowering of the HOMO of the metal complexes, this could be caused by changes in the electronic state of the complex as it stacks with neighbouring bases. For **CuL** the observed band in Figure 4.9b became evident as a shoulder of a larger band at 290 nm. The nickel species appears more strongly coloured (pink/red) than the copper analogue (yellow).

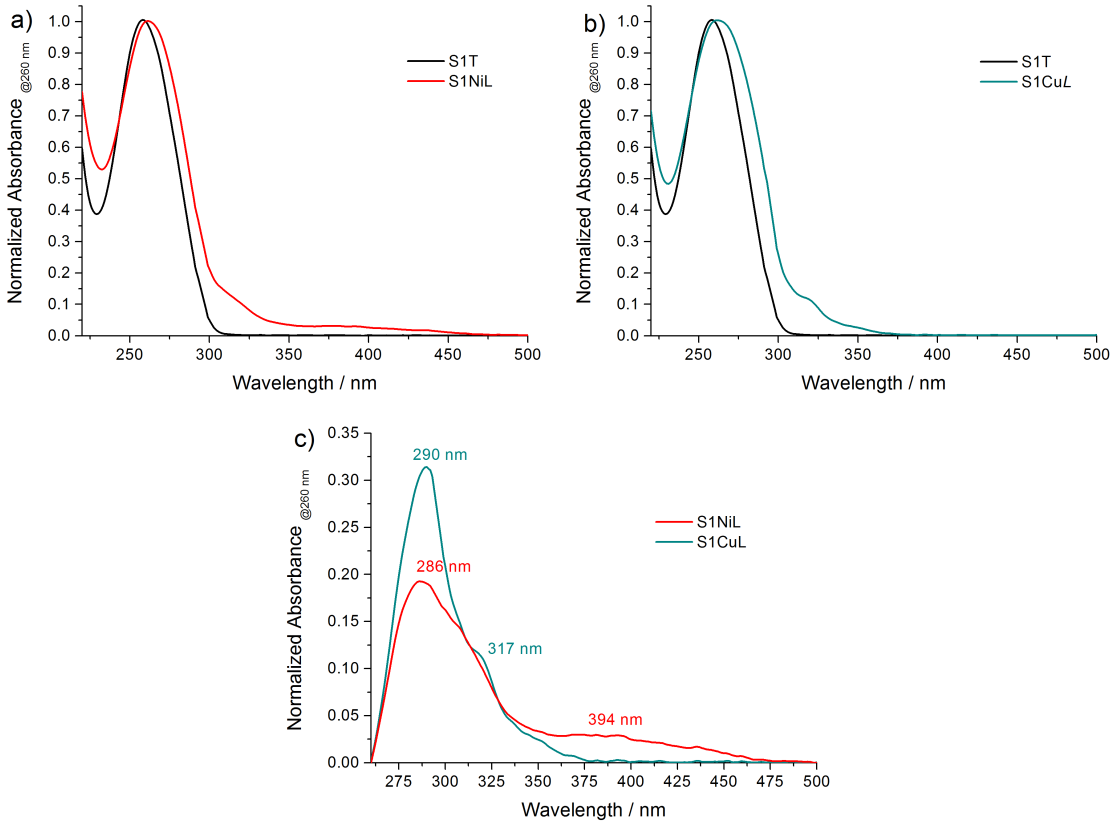


*Figure 5.12 - Normalised (@ 260 nm) UV-vis spectra of a) **S1NiL** and b) **S1CuL**. Unmodified sequence S1T is shown for comparison. c) Shows the main macrocycle contributions following subtraction of the DNA spectra for **S1Ab**. 10 mM sodium phosphate buffer, 100 mM NaCl.*

Incorporation of multiple **CuL** or **NiL** into the backbone of **S1** or **S2** has varying effects on the electronic spectra and there is some sequence dependence on the appearance and intensity of the bands. Figure 5.13a compares one incorporation of **NiL** *versus* two in two different strands **S1NiL**, **S1NiLNiL** and **S2NiLNiL**, This analysis shows a significantly enhanced Soret band at 320 nm for the doubly incorporated systems, which is also considerably shifted (286 – 320 nm) by virtue of neighbouring nucleobases (C**NiLNiL**C in **S1** *vs.* G**NiLNiL**G in **S2**) or perhaps the close proximity of neighbouring macrocycles. Bismacrocyclic complexes of this type of compound have been shown previously to be able to interact significantly with each other to induce changes in electron donating behaviour.[60,61]

Similarly **CuL** shows some sequence dependence in the intensity of its strong absorption, at 290 nm (Figure 5.13b).

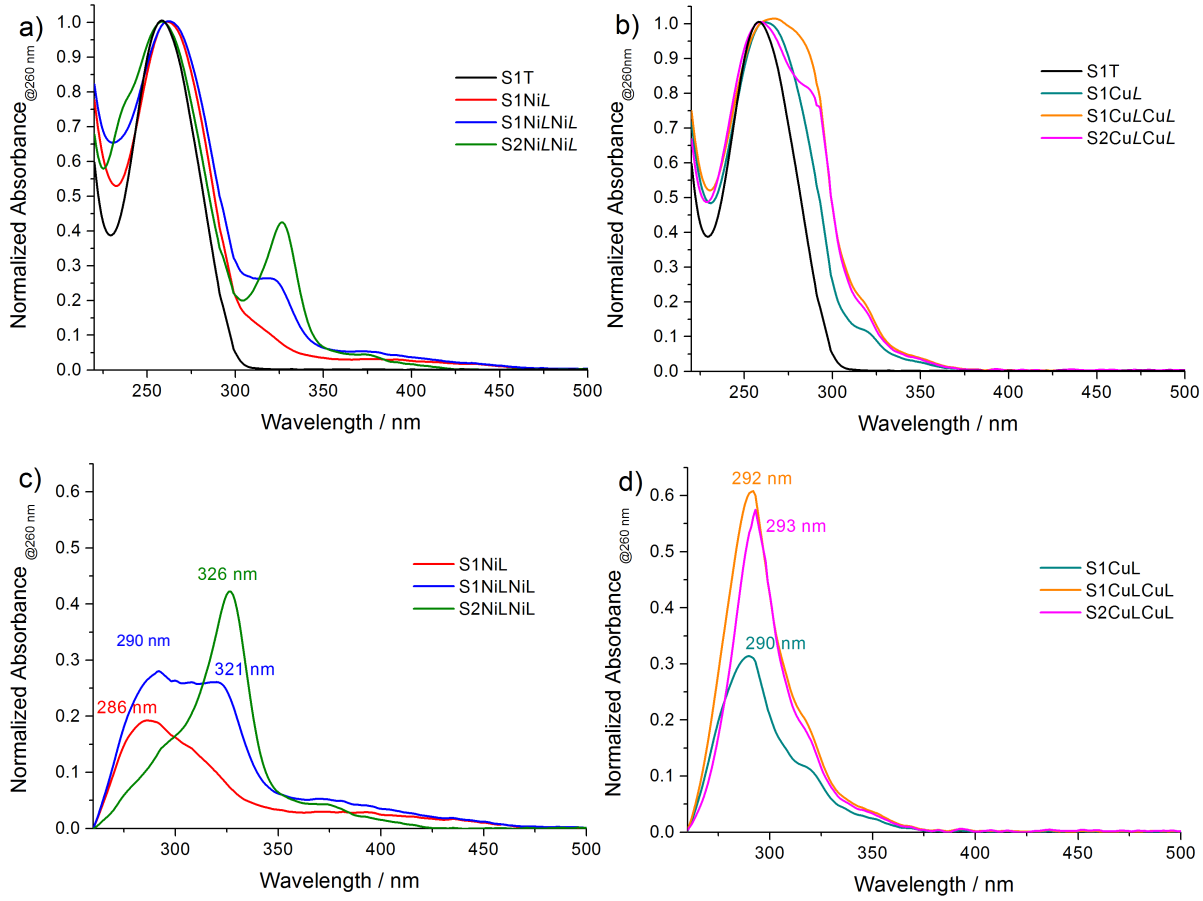


*Figure 5.13 - Normalised UV-vis spectra of a) Strands **S1** and **S2** labelled with multiple **NiL**. b) Strands **S1** and **S2** labelled with multiple **CuL** and macrocycle contributions following subtraction of **S1Ab** for both Ni (c) and Cu (d) incorporations.*

5.2.4.2 Thermal Melting Studies

Analysis of the thermal melting profiles of the linker complexes **CuL** and **NiL** was carried out using variable temperature UV-vis spectroscopy to establish the effect of the bridging ligand on duplex stability. For a more detailed explanation of thermal melting analysis of DNA see Chapter 2, section 2.2.1. Duplexes **S1X•S2Y** (where X = **CuL/NiL**, Y = A or AA) showed a considerably diminished stability (>16 °C) over fully complementary sequence **S1T/TT•S2A/AA** by virtue of the observed T_m values (see Table 5.2). This was not entirely unexpected given that incorporation of the linker complexes involves the loss of stabilising T-

A base pairs. The reduction in T_m for these modifications suggests that no stacking (intercalation) is observed for these linker complexes, as expected from the design rationale (Scheme 5.1b).

*Table 5.2 - Melting temperatures (T_m) of duplexes **S1X•S2Y** (X= CuL/NiL, Y = A/AA). Unmodified sequences where X = T/TT and Y = A/AA are shown as a comparison. 5 μM of each strand in 10 mM sodium phosphate buffer, 100 mM NaCl.*

	T_m / °C	
S1 X =	**S2A**	**S2AA**
T	55.0 (± 0.0)	--
TT	--	56.0 (± 0.2)
Cu*L*	38.5 (± 0.6)	40.0 (± 0.3)
Ni*L*	39.0 (± 0.0)	39.0 (± 0.3)

(.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

Multiply labelled systems show positive stabilising contributions (see Table 5.3) for the linker complexes, perhaps afforded by strong pi-pi interactions between adjacent complexes in the duplex. Having the same number of incorporations in each strand is evidently important to the overall structure and stability, ensuring that the nucleobases are spaced out evenly within the duplex and can still participate in hydrogen bonding motifs. Seemingly the metal itself plays a very small part in those interactions given that **S1CuLCuL•S2CuLCuL** and **S1CuLCuL•S2NiLNiL** have the same T_m. The reverse sequences only differ by 1 °C.

Table 5.3 - T_m values of duplexes bearing multiple linker modifications. 5 μM of each strand in 10 mM sodium phosphate buffer, 100 mM NaCl.

	T_m / °C*	
	S2 Y =	
S1 X =	**Y = CuLCuL**	**Y = NiLNiL**
CuL	44.0 (± 0.1)	--
CuLCuL	49.0 (± 0.2)	49.0 (± 0.0)
NiL	--	45.0 (± 0.6)
NiLNiL	49.0 (± 0.5)	50.0 (± 0.2)

(.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

5.2.4.3 Circular Dichroism Studies

Circular dichroism spectra of these complexes in the previously described duplex **S1X•S2AA** at room temperature are given in Figure 5.14. They show small differences despite the large destabilisation in the T_m. However the helicity and hence CD bands that are indicative of B-DNA (see Chapter 2) structure are slightly altered by the presence of the linker when compared with the unmodified duplex **S1TT•S2AA**. The copper complex (T_m = 40 °C) shows a slightly enhanced B-DNA structure, perhaps an indication of the increased π-rich surface area enforced by the planar, aromatic nature of **CuL** which is due to the larger metal ion centre. The size of the copper macrocycle has a larger effect on the stacking interactions as shown by the differences in T_m when the cavity is increased in size (**S2A** to **S2AA**). A small negative band appears for both the Cu and Ni complexes at approx. 290 nm, which has previously been ascribed to the 'flipping out' of an adjacent nucleobase in different systems (Chapter 3). However, the origin of this band is likely to be due to an induced CD (ICD) band from the complexes themselves upon incorporation into a helical DNA backbone. The larger intensity for the copper modified **S1** is as a result of an increase in extinction coefficient at 290 nm (see Figure 5.12).

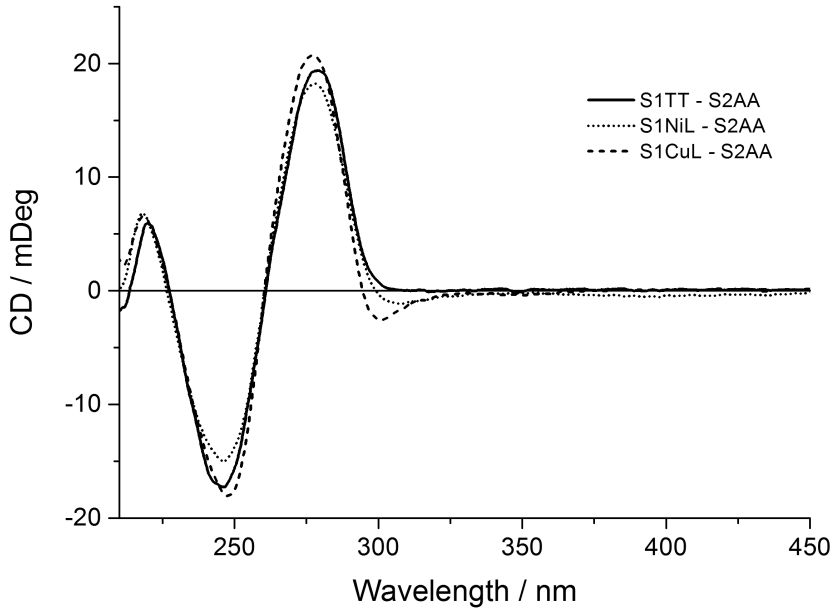


*Figure 5.14 - CD spectra of **S1NiL/CuL•S2AA**. Unmodified sequence **S1TT•S2AA** is shown for comparison. 5 μM of each strand in 10 mM sodium phosphate buffer, 100 mM NaCl.*

5.2.4.4 Electrochemical behaviour

Similar tetrazacyclene macrocycles (Figure 5.15) have been previously shown to possess *quasi*-reversible electrochemical behaviour,[59,61] where the M(II) cation undergoes a 1-electron oxidation to the 3+ oxidation state. The nickel complex exhibits an oxidation potential (*vs.* Ag/AgCl) of 783 mV, 217 mV positive of that of the comparable copper complex (566 mV). The nickel macrocycles exhibit the higher oxidation potential due to the instability of Ni(III) and the resulting distortion of the square planar geometry (d^8 to d^7) of the macrocycle. In comparison the copper macrocycles undergo a more favourable oxidation to Cu(III). Oxidation to a d^8 metal centre would be facilitated by the square planar arrangement of donors about the macrocyclic ring and would not require significant reorganisation of the macrocycle.[59]

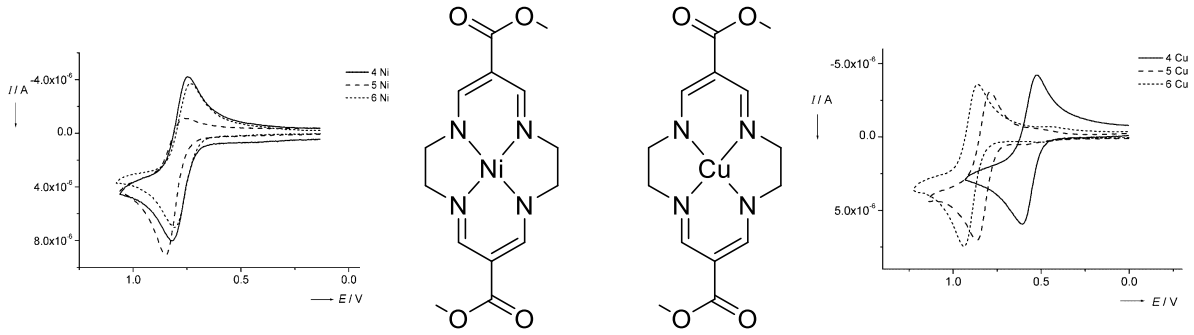


Figure 5.15 - Nickel and copper macrocycles previously reported to exhibit quasi-reversible electrochemical behavior. The complexes shown are depicted by the solid line in the cyclic voltammograms. The potentials are vs. Ag/AgCl. Figures adapted from ref. 59.

An investigation of the linker-modified strands (**S1NiL and S1CuL**) for their electrochemical potential was carried out using cyclic voltammetry and square wave voltammetry. For a more in-depth discussion of these techniques see Chapter 2. The samples consisted of 50 μM solution of the strand in a 100 mM Tris-HCl buffer (pH 7.0) and 10 mM NaCl solution.

Whilst the required working concentrations for studying these complexes by electrochemistry was well above (orders of magnitude) that required to be useful in a potential device, the study at these concentrations enables the understanding of the behaviour of these complexes. Future work would require surface immobilisation or optimisation of electrochemical techniques in order to reduce both det ection limits and material requirements.

S1CuL exhibits reversible ($\Delta E_p \sim 60$ mV) electrochemical behaviour with a potential ($E_{1/2}$) for the one electron oxidation of Cu(II) to Cu(III) at 423 mV when studied with cyclic voltammetry (Figure 5.16). The reversibility is also highlighted by Randles-Sevćik plots (Figure 5.16b) which display a linear relationship between the observed currents and square root of the scan rate.[63] The redox process can also be monitored with square wave

voltammetry (Figure 5.16c) and a $E_{1/2}$ value of 428 mV is obtained. The more negative oxidation potential of these macrocycles in comparison with those previously reported (566 mV)[59] is likely due to contributions of the surrounding π systems and cation stabilisation by negative phosphates of the DNA backbone.

The electrochemical behaviour of **S1CuL** upon addition of **S2AA** was monitored with square wave voltammetry (Figure 5.16c), in which increasing amounts of the target strand, and hence the formation of a duplex, resulted in a decrease in the observed current of 33% upon full hybridisation. These changes are not thought to be a result of larger changes in the linker's local environment but of changes in the bulk properties of the duplex itself. An increase in size and change in overall charge is probably responsible for the observed decrease. Similar changes were observed for different targets as a result the linker system is not suitable as a SNP sensor.

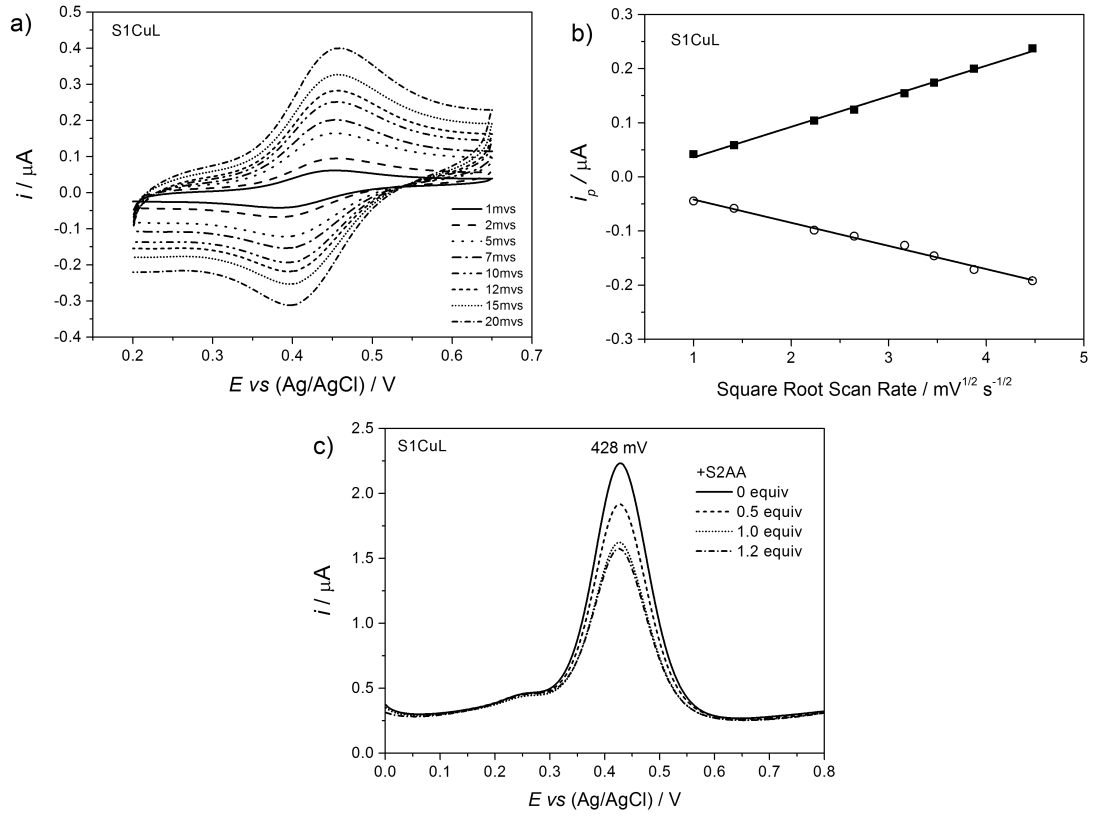


*Figure 5.16 - a) Scan rate behaviour of Cu linker **S1CuL** and b) Randles-Sevick plots of the scan rate dependence for **S1CuL**. ■ – i_p^a, ○ - i_p^c. c) Square wave voltammograms of **S1CuL** upon addition of **S2AA**. 50 μM of target strand in 10 mM sodium phosphate buffer, 100 mM NaCl.*

*Table 5.4 - Electrochemical behaviour of **S1CuL** as measured by CV (15 mv s^{-1}). 50 μM, in 10 mM Tris-HCl buffer pH 7.0, 100 mM NaCl.*

Strand	**CV**				**SWV**
	E_p^a / V	E_p^c / V	E_p^a-E_p^c / V	$E^{o'}$ / V	E_p / V
S1CuL	0.452	0.392	0.062	0.423	0.429

In contrast to the copper system, the nickel complex **S1NiL**, displays no observable redox activity in the potential window of -1 V to +1 V (*vs*. Ag/AgCl). Whilst not fully expected, this lack of activity could be attributed to the increased constraints in distortion of the square planar geometry, imparted on the macrocycle by the stacking interactions of the neighbouring and opposing nucleobases. For aqueous systems the expected oxidation

potential of the nickel complex, approx. 200 mV greater than that of the copper, means observing currents is hindered by the oxidation and absorption of buffers and salts in the electrochemical matrix.

5.2.5 Intercalating Systems (Single Strands)

5.2.5.1 UV-vis Characterisation

UV-vis investigation of the intercalating modifications (**Cu*a*, Cu*b*, Ni*a* and Ni*b***) to **S1** provided a wealth of information about the electronic and structural properties of these moieties. The incorporation of these compounds with their electronic properties is significant enough that the sequences appear coloured to the eye (**S1Ni*a/b*** = pink/red, **S1Cu*a/b*** = yellow) at μM concentrations. Both nickel and copper-centred macrocycles exhibit Soret-like bands like those previously described. Figure 5.17a)b) shows the normalised UV-vis spectra obtained from the singularly modified **S1X** sequence where X = Ni/Cu. The nickel complexes (*a/b)* have a distinct band at 326 nm while the analogous copper complexes display what appears to be a smaller band at 317 nm. However, upon further analysis (Figure 5.17d) and subtraction of the spectra of the abasic sequence (**S1Ab**), it became apparent that a larger Soret band at 290 nm was present, followed by a shoulder at 317 nm. These observations mimic, albeit at higher energies, the transitions observed for the similar designs described by Piela *et al*.[61]

In the case of both metal centres, the slow isomer, ***b***, displays a slight reduction in the intensity of the macrocyclic transitions, along with a small red shift (1-2 nm). This suggests that even in the single stranded form, the macrocycles interact and perhaps stack differently with the flanking nucleobases.

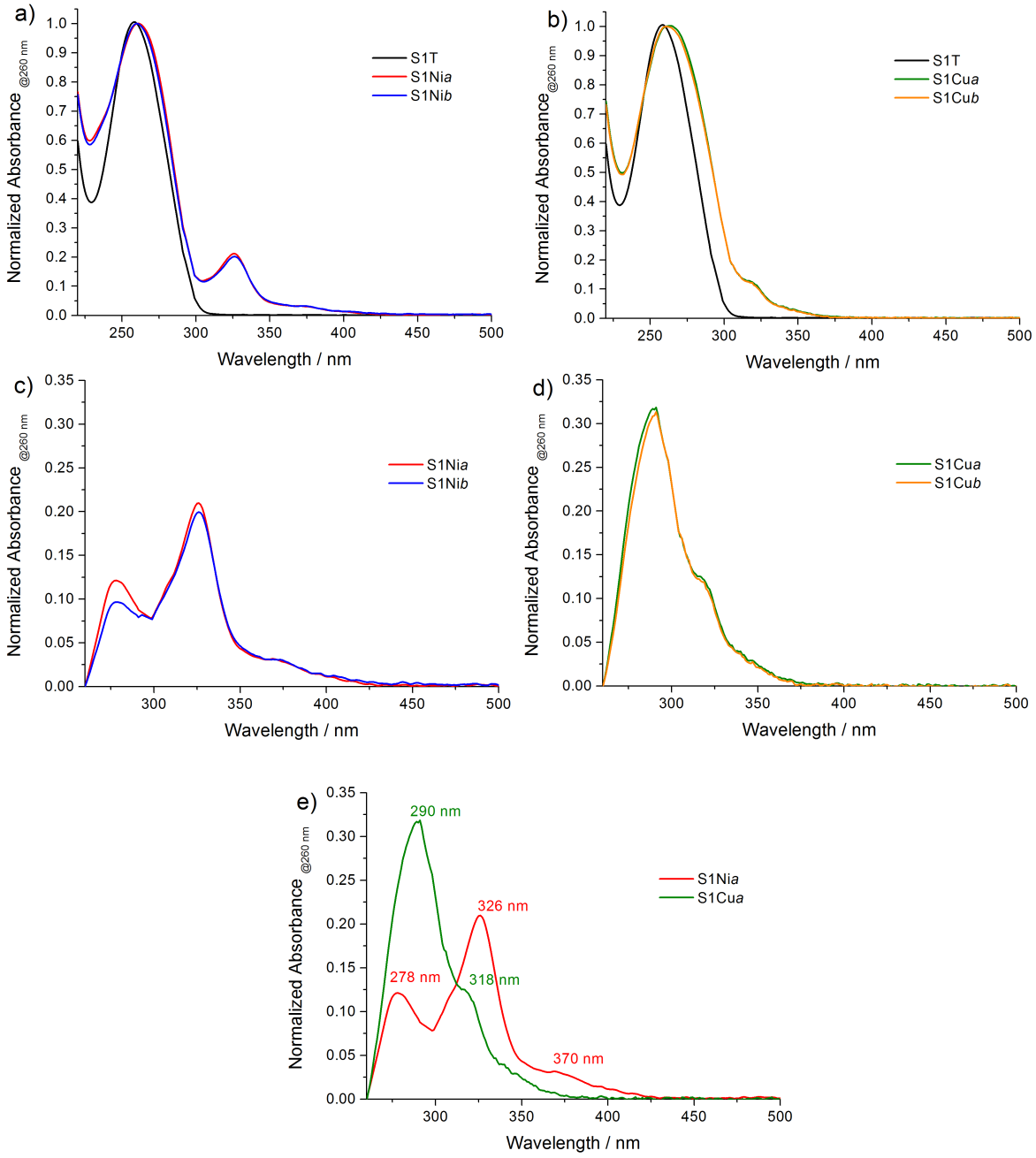


*Figure 5.17 - Normalised (@260 nm) UV-vis spectra of a) **S1Nia** and **S1Nib** compared with unmodified **S1T**. b) Normalised UV-vis spectra zoomed in on the MLCT bands of **S1Nia** and **S1Nib**. UV-Vis spectra of showing the contributions of the macrocycles following subtraction of normalized DNA spectra (**S1Ab**) for opposing isomers of both Ni (c) and Cu (d) complexes. e) Overlaid UV-vis spectra comparing the macrocyclic contributions of both complexes following subtraction of normalized DNA spectra (**S1Ab**). 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.*

5.2.5.2 Electrochemical Behaviour

Electrochemical interrogation of the intercalating derivatives of the modified sequences **S1Cua** and **S1Cu*b*** again showed reversible electrochemical behaviour (Figure 5.18), as evidenced by ΔE_p ~ 60 mV and the observed currents are linearly proportional to the square root of the scan rate. The strands display $E_{1/2}$ potentials of 439 and 444 mV for the fast and slow isomers respectively. Their differing potentials likely result from their respective geometries and hence orientation in the nucleobase stack, the contributions of which affect, albeit by a small amount, the electron donation abilities of the complex.

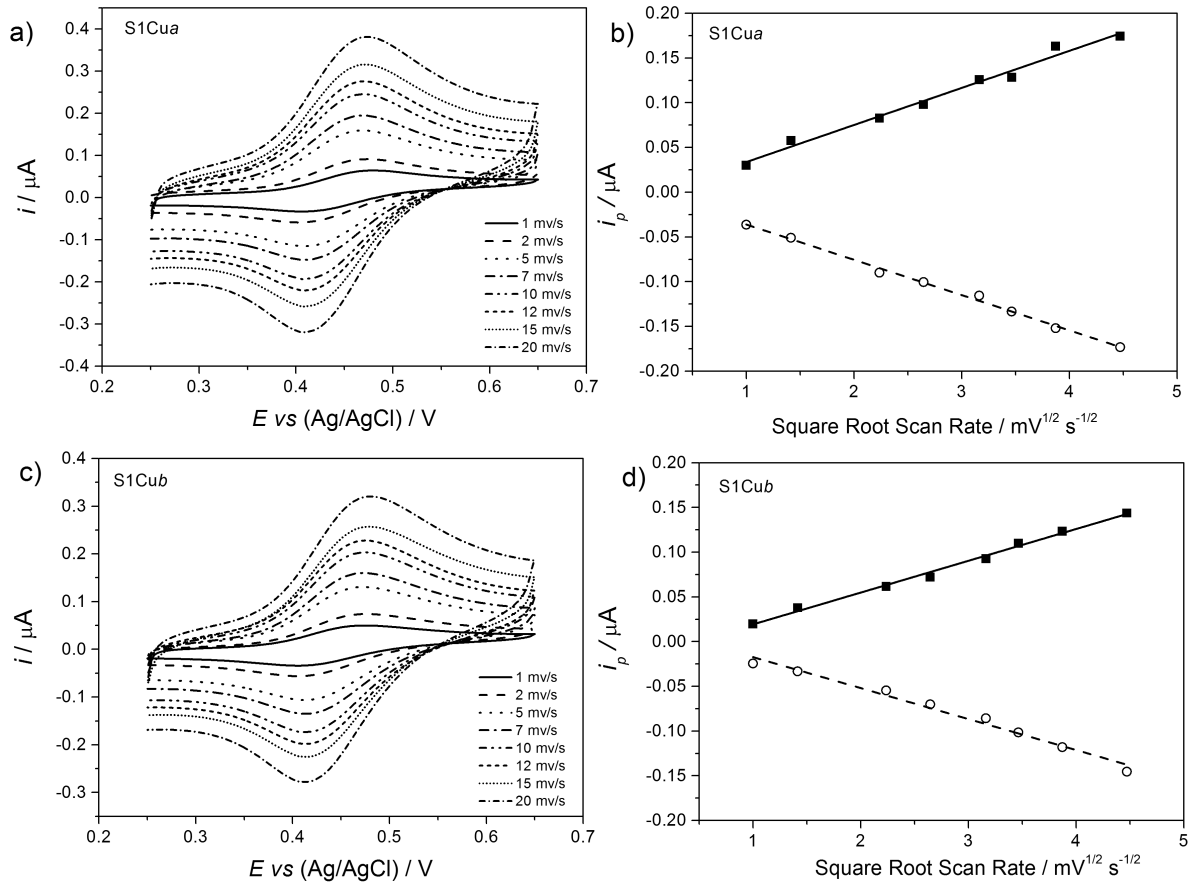


*Figure 5.18 - Scan rate behaviour of Cu intercalators a) **S1Cua** b) **S1Cub** and plots of i_p vs. $v^{1/2}$ for c) **S1Cua** and d) **S1Cub**. ■– i_p^a, ○ - i_p^c.50 μM DNA, 100 mM NaCl, 10 mM Tris-HCl pH 7.0.*

Square wave voltammograms of the single strands, given in Figure 5.19, show similar differences in the observed $E_{1/2}$ potentials. Table 5.5 below details the measured redox potentials of these systems.

*Table 5.5 - Redox properties of the **Cua** and **Cub** complexes when incorporated into S1 (15 mv s^{-1}). 50 μM, in 10 mM Tris-HCl buffer pH 7.0, 100 mM NaCl.*

Strand	CV				SWV
	E_p^a / V	E_p^c / V	E_p^a-E_p^c / V	$E_{1/2}$ / V	$E_{1/2}$ / V
S1Cu*a*	0.471	0.407	0.062	0.439	0.434
S1Cu*b*	0.477	0.411	0.064	0.444	0.440

Values are described as the average of a minimum of three repeats. Error associated with potentials is ±3 mV throughout. Experiments carried out at 50 μM of S1, 100 mM NaCl, 10 mM Tris-HCl at room temperature.

The background suppression characteristics of SWV make it a desirable technique for biosensing. Notably, however, the voltammograms shown in Figure 5.19 show a defined pre-peak, which was not observed in the CVs shown in Figure 5.18. Pre-peaks are related to the adsorption of electrochemical products to the electrode and are more common with glassy carbon electrodes.[63] This pre-peak will be further discussed in 5.3.5.5.

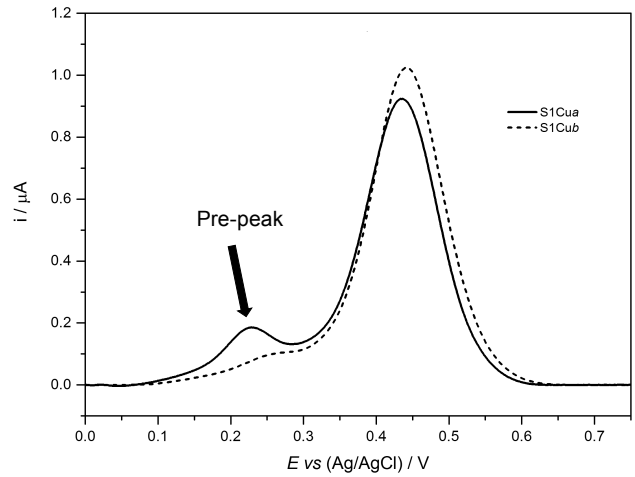


*Figure 5.19 - Square Wave Voltammograms comparing **Cua** and **Cub** when incorporated into **S1**. Solid line - **S1Cua**. Dashed line - **S1Cub**. 50 μM DNA, 100 mM NaCl, 10 mM Tris-HCl pH 7.0.*

In contrast, the nickel motifs displayed no observable redox process as also found for the linker systems. Therefore although the nickel complexes could not be used for electrochemical sensing they remain interesting from both a structural and electronic point of view and indeed give more defined Soret bands as shown in Figure 5.17.

5.2.6 Intercalating Systems (Duplex Studies)

5.2.6.1 UV-vis Studies

UV-vis inspection of the macrocyclic complexes upon binding of DNA target **S2A** provided insight into the behaviour of the systems upon duplex formation. Figure 5.20 shows an example of how the characteristic Soret band (326 nm) of the nickel complex is red-shifted upon formation of a duplex with **S2A** (comparisons of the two isomers can be found in appendix 13). Red-shifting of absorbance bands is a characteristic feature of intercalation processes.[2,64,65] The exoergonic process of DNA binding and interaction with the base-stack results in a lowering of the HOMO of the bound chromophore and hence causes a red shift, while decreases in the intensity (hypochromicity) suggests π interactions with the nucleobases. The copper complexes exhibit a similar red shift, although they are harder to interpret given the overlapping DNA signatures.

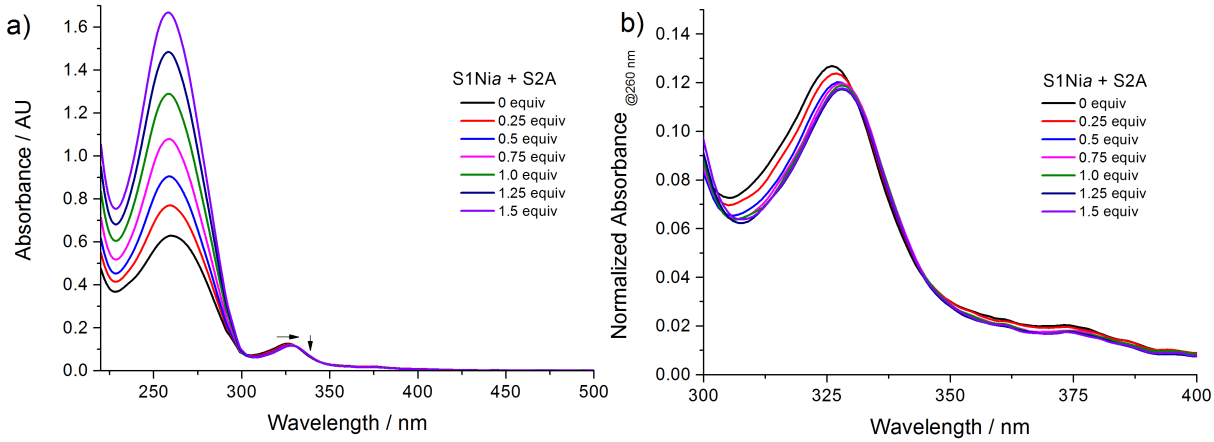


*Figure 5.20 - a) UV-vis titration of **S2A** to **S1Nia** showing the b) red shift in Soret band upon binding. 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.*

5.2.6.2 Thermal Melting Studies

Thermal melt analys*i*s of **S1M*a/b*** systems when bound to sequence **S2A**, which bears an opposing adenine, helped provide further insight into the ability of each isomer to intercalate. Considerable differences were observed (see Table 5.6) through simply changing the stereochemistry with the T_m values of the nickel complexes differing by 5 °C and the fast isomer, **S1Ni*a*** displaying the highest stability. Similarly, the copper analogues show the same higher stability for the fast isomer (**S1Cu*a***), albeit of a smaller magnitude (2.5 °C). Comparison of the T_m values with that of unmodified sequence **S1T•S2A** (T_m = 55 °C) shows that the intercalators, despite a loss of hydrogen bonding are able to stabilise the duplex presumably through π-π interactions between the complexes and the parallel nucleobases (i.e. intercalation). Consideration of the electronic properties of the two metal centres provided some rationale for the differences afforded by the two metals. The nickel macrocycles are less electron-rich than their copper partners, a factor that would lead to enhanced affinity for π-donors, hence stronger stacking interactions with the neighbouring nucleobases. Differences caused by the change in stereochemistry (***a*** *vs* ***b***) are likely due to the change in approach the intercalator must make to 'slot' into the duplex cavity.

*Table 5.6 - T_m results for **S1Ma/b** (where M = Cu/Ni) when duplexed with **S2A**. 5 μM of each strand, in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl. (J-LD)*

S1 X =	T_m / °C* **S2 Y = A**	ΔT_m / °C
T	55.0 (± 0.0)	-
Ni*a*	55.0 (± 0.2)	0.0
Ni*b*	50.0 (± 0.9)	-5.0
Cu*a*	52.5 (± 0.3)	-2.5
Cu*b*	50.0 (± 0.5)	-5.0

Δ*T*m relative to unmodified sequence S1T•S2A (55 °C). (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

5.2.6.3 Molecular Models

Molecular modelling studies (carried out by J-L.D using Insight II) (Figure 5.21), shows the predicted structures for **S1Cu*a*** and **S1Cu*b*,** based on *S* and *R* stereochemistry respectively when bound to **S2A**. These depict the macrocycle approaching intercalation through two very different angles. The fast isomer, **S1Cu*a*,** follows a left hand twist, the opposite of that of the B-DNA helix and inserts *via* the major groove. This orientation suggests a limited amount of disruption to base stacking and the helical twist of DNA, where the macrocycle slots below the plane of the opposing nucleobase and stacks effectively with neighbouring bases. The slow isomer, **S1Cu*b***, in comparison, causes a more significant disruption with the opposing nucleobase flipping out as it inserts *via* the minor groove. Both clearly show insertion into the hydrophobic DNA core, stacking and interacting closely with surrounding nucleobases. The added disruption afforded by **Cu*b*** is in line with the lower T_m values observed. Furthermore, study of the CD spectra (Figure 5.22) shows the stark differences in the overall structure of the copper macrocycle modified duplexes.

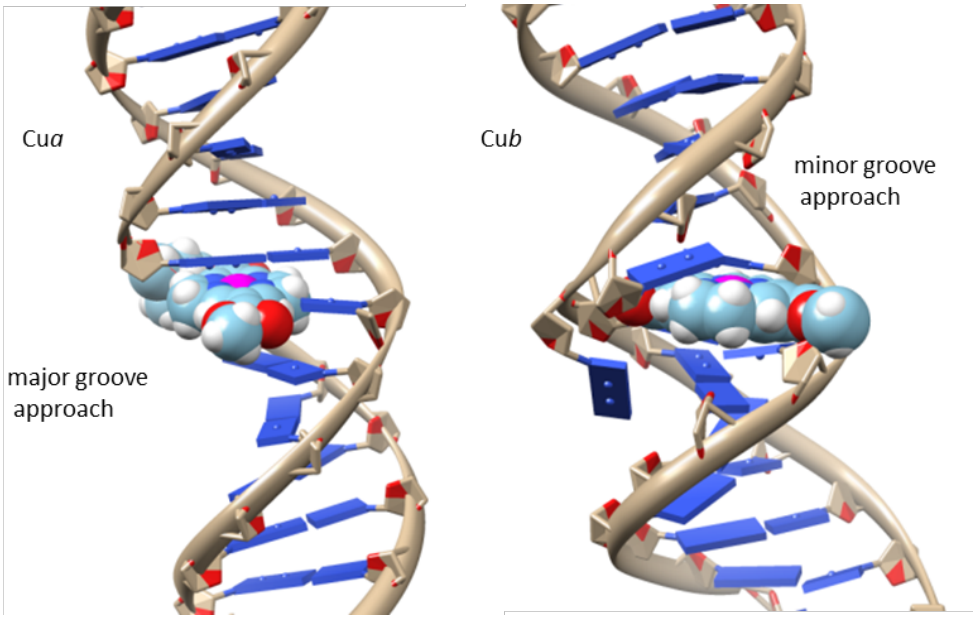


*Figure 5.21 - Molecular models of the two stereoisomers of **S1Cua** and **S1Cub** indicating the different approaches to intercalation when bound to **S2A**. (Models created by J-L.D using Insight II).*

5.2.6.4 Circular Dichroism Studies

Both macrocycles display an ICD for the intercalation of the macrocycle into the helical twist of the duplex, characterised by three distinct bands, all of which are observed in the UV-vis spectrum (Figure 5.17). Notably the small shoulder observed at approx. 360 nm is relatively large in comparison with the other bands, which indicates that this particular transition is orientated more in line with the helical twist of DNA. The characteristic B-DNA signals, evidenced by a negative band at 245 nm and positive band at 275 nm, are significantly reduced as a result of incorporation and intercalation of both isomers of the macrocycle, which suggests an unwinding or disruption of the DNA base stack. Similarly to the T_m results, the two isomers display marked differences in their respective spectra. The fast isomer, *a*, has a slightly reduced B-DNA structure which supports the left handed stereochemistry of the postulated structure and molecular simulations. Conversely, the fast isomer is characterised by an enhanced negative band at 290 nm, that of the main Soret band of the copper complex. It is believed that this difference is caused by a 'flipping out' of an opposing or surrounding nucleobase, its changed orientation affecting the overall intensity of the band. The other macrocyclic bands are consistent in position and intensity for both isomers which helped for the postulation of this structural change, which is reinforced by the molecular dynamic simulations.

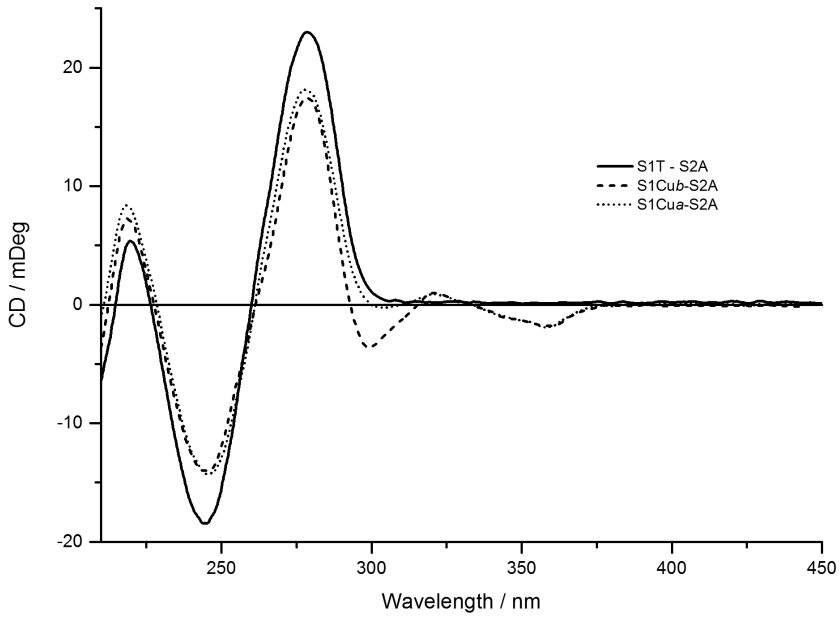


*Figure 5.22 - CD spectra of duplex **S1Cua/b•S2A**. Unmodified control **S1T•S2A** is also shown.*

5.2.6.5 Varying the Base Opposite

For a SNP sensor to be effective, ideally the local environment of the tag should vary depending on the nucleobase in its proximity, which could bring about differences that are a result of factors such as changes in cavity size, redox potential or stacking motifs. After evidence for the intercalative binding mode was established, the nucleobase opposite in the target sequence was varied. Given the electrochemical properties of the single strand, it was proposed that the base opposite could influence the electrochemical output of the redox active copper macrocycles and this became the focus of the investigation. Thermal melting analysis (Table 5.7) of **S1Cua/b•S2Y** (Y = A, G, C, T) established contrasting differences between the overall duplex stabilities. Purines, adenine and guanine, being of similar size, show less stability over pyrimidines, thymine and cytosine when opposite the macrocycle in the sequence. When a pyrimidine (C/T) lies opposite the fast isomer the T_m is even greater than that of a fully matched **S1A•S2T** (55 °C) duplex, suggesting the macrocycle offers enhanced π-π interactions over adenine. This is not unsurprising, given its larger surface

area and electron deficient structure. The interactions may also be enhanced by metal-nucleobase interactions between the copper and the electron rich donors.

*Table 5.7 - T_m results when **S1Cua/b** is bound to **S2Y** with varying bases opposite. 5 μM, in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl. (J-LD and JC-S)*

S2Y	T_m / °C* (ΔT_m / °C)# **S1X =**			
Y =	**Cu*a***	ΔT_m / °C#	**Cu*b***	ΔT_m / °C#
A	52.5 (± 0.3)	-2.5	50.0 (± 0.5)	-5.0
G	54.5 (± 0.0)	-0.5	50.0 (± 0.0)	-5.0
C	57.0 (± 0.0)	+2.0	54.0 (± 0.0)	-1.0
T	57.0 (± 0.0)	+2.0	55.0 (± 0.2)	-
Abasic	57.0 (± 0.2)	+2.0	58.5 (± 0.2)	+3.5

#ΔTm relative to unmodified sequence **S1T•S2A** (55 °C). # Δ*T*m relative to unmodified sequence S1T•S2A (55 °C). (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

Reductions in the T_m in the presence of an opposing purine could result from the reduced cavity size afforded by the larger nucleobases, which result in a distortion of the DNA backbone; in order to maximise stacking interactions and with the macrocycle, the base stack of the duplex would have to be distorted. Figure 5.23 below shows a schematic representation of the possible effect of varying the size of the base opposite the macrocycle, when opposite thymine, the macrocycle could still participate in additional favourable stacking interactions with flanking guanines in the target sequence.

A significant finding was that complete removal of the opposing nucleobase, creating an abasic site, results in the highest T_m, which is 2-2.5 °C higher than the unmodified sequence. The T_m values of these strands provide further evidence for the directionality of the stereochemistry of the two macrocycles. The proposed *R* isomer, following the same right hand twist of B-DNA when unrestricted in its intercalation pathway, has the highest observed T_m of 58.5 °C. In contrast the *S* isomer, having left handed stereochemistry, implements some distortion in the DNA backbone and hence the duplex stability is 2.5 °C lower.

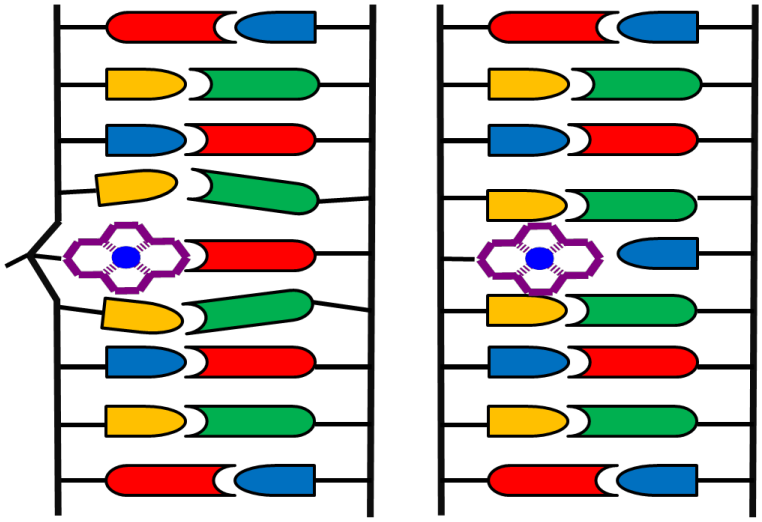


Figure 5.23 - Schematic representation of the effects of the base opposite Cua/b in S1, where S2Y = purine(A) or pyrimidine (T).

5.2.6.6 Multiple Incorporations

Incorporation of multiple macrocycles into the duplex has the effect of enhancing the overall stability when the number of incorporations in the duplex increases to two (one opposite the other). However its effect is dependent on the stereochemistry (Table 5.8). Increased incorporations to three (**S1Cua*/b*•S1Cua*/b*Cua*/b***) and four macrocycles (**S1Cua*/b*Cua*/b*•S1Cua*/b*Cua*/b***) leads to loss of stability, but seemingly with these larger numbers, having opposing stereochemistries plays some part in alleviating the additional helical strain of incorporating non-natural nucleosidic mimics into the duplex. While the study of these 'zipper-like' systems[66,67] is interesting to the groups ongoing understanding of this new category of DNA modifications, their use was not thought to be conducive to SNP sensing, and they were not studied further in this project.

*Table 5.8 - T_m results when multiple incorporations of copper macrocycles (**Cua/b**) are placed in **S1** and **S2**. 5 μM, in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl. (J-LD)*

S1 X =	T_m / °C **S2 Y =**							
	Y = Cu*a*	ΔT_m / °C#	**Y = Cu*b***	ΔT_m / °C#	**Y = Cu*a*Cu*a***	ΔT_m / °C#	**Y = Cu*b*Cu*b***	ΔT_m / °C#
Cu*a*	57.5 (± 0.7)	+2.5	60.5 (± 0.7)	+5.5	53.0 (± 0.2)	-2.0	54.0 (± 0.2)	-1.0
Cu*b*	55.0 (± 0.6)	0.0	59.0 (± 0.2)	+4.0	55.0 (± 0.4)	0.0	54.0 (± 0.2)	-1.0
Cu*a*Cu*a*	-	-	-	-	53.0 (± 0.4)	-2.0	54.5 (± 0.3)	-0.5
Cu*b*Cu*b*	-	-	-	-	52 .0 (± 0.2)	-3.0	52.0 (± 0.3)	-3.0

ΔT_m relative to unmodified sequence **S1T•S2A** (55 °C). (.) shows the standard error of the mean to the nearest 0.1 °C. Error in instrument data = ±0.5 °C.

5.2.6.7 Electrochemical Sensing

Such is the design of these intercalating modifications that it was important to investigate whether the nature of the intercalation behaviour was conducive to differences in electrochemical output as function of the sequence and, more specifically, the base opposite. The use of square wave voltammetry allowed for the more accurate measurement of the output of the copper macrocycle given that the technique removes much of the background charging currents (for a more detailed discussion on analytical electrochemical techniques see Chapter 2). Initial measurement of the single stranded oligonucleotides (**S1Cu*a* + S1Cu*b***) was followed by additions of target sequence **S2T**, with the linker motif, a reduction in the current was observed (Figure 5.25b+d). For the fast isomer, a reduction in the current of 28% was observed; similarly, for the slow isomer a 25% decrease was seen. Whilst it was predicted that large changes in bulk properties were the main cause of such reductions, as previously seen with FcNA modifications (Chapter 2), it was hoped that the more penetrative insertion of the macrocycle and locus of the metal centre in the DNA core could bring about measurable changes. It was noticeable that the pre-peak observed for both complexes displays no change in current intensity as the target is added (Figure 5.24). This observation is consistent with a believed adsorbed product of the oxidation being present at the surface;[63]

because of its weak immobilisation, it is unable to partake in binding to the target and hence its current contribution remains constant.

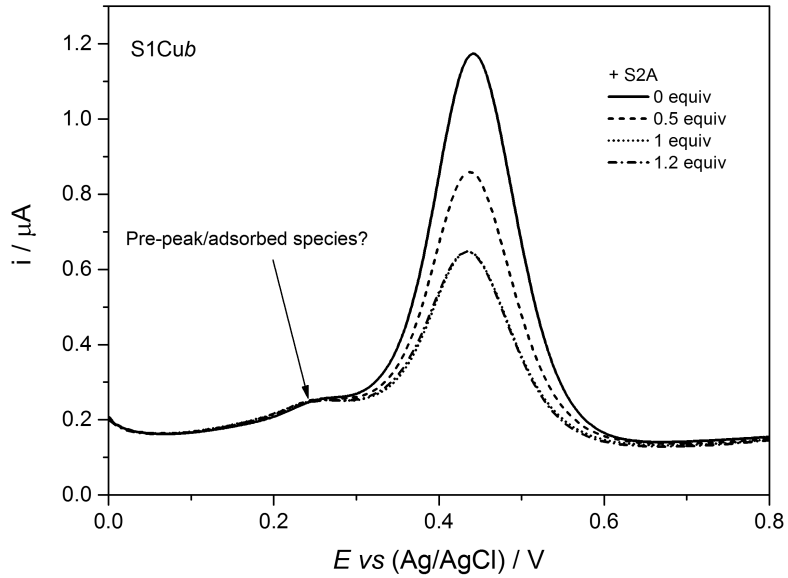


*Figure 5.24 - Example of unedited SWV for the addition of **S2A** target to **S1Cub**. 50 μM, in 10 mM Tris-HCl buffer pH 7.0, 100 mM NaCl.*

The single strands were then exposed to **S2A** to form the corresponding duplexes. The differences observed for each isomer were considerable, with the fast isomer showing a similar current decrease to that when bound to **S2T** (29%). However when **S1Cu*b*** was bound to **S2A**, the decrease in current was markedly different with a two-fold reduction of 52% (Figure 5.25), to indicate the electrochemical sensing of a single base change in a sequence of target DNA.

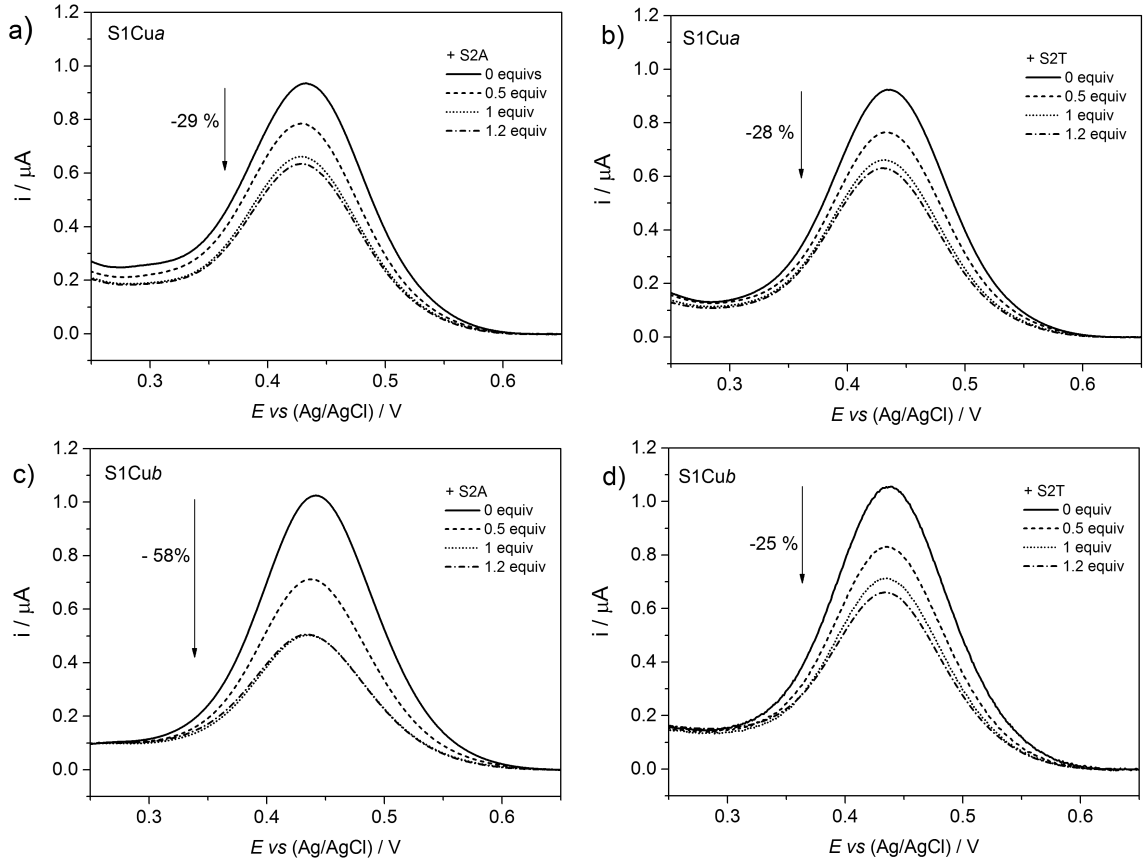


*Figure 5.25 - Square wave voltammetry of the initial screens of **S1Cua/b** when targets **S2A** or **S2T** were added. 50 µM of **S1** in 10 mM Tris-HCl buffer pH 7.0, 100 mM NaCl.*

The ability to discriminate between adenine and thymine is believed to be due mainly to size contributions of the opposing nucleobases themselves. Given that **S1Cu*a*** was unable to discriminate between purine and pyrimidine, the focus of the SNP sensing was driven towards full investigation of **S1Cu*b***. The ability of one isomer to be more effective than the other is not unexpected, as this has been found previously with fluorescence-based anthracene SNP probes (see 5.1.2.1).[8] Additionally the molecular models and circular dichroism results, suggest that the two stereoisomers interact with the duplex differently, although both are intercalating. According to the models the long axis is perpendicular to the opposing nucleobase for **Cu*a*** but aligned for **Cu*b***. It is possible that this is a factor, which leads to the **Cu*b*** being able to feel the effects of minute changes more markedly.

Probe **S1Cu*b*** was then investigated further with other canonical nucleobases guanine and cytosine and the effects monitored again by SWV (Figure 4.25). Remarkably the macrocycle was shown to be able to discriminate between all four bases *via* reduction in current intensities; although the discrimination between G (-47%) and C (-39%) was not as prominent these changes are outside of experimental error (±1.3% (S2G), ±2.5% (S2C) (Figure 5.26 and Table 5.9). Importantly, and perhaps beneficial to the future development of assays based on this system, the discrimination between T and A is considerable. T to A base changes have been shown to be particularly prominent in certain cancer strains and therefore these probes could well be suitable for investigating SNP sites in known genomic sequences e.g. BRAF system where a T to A base variation accounts for 92% of the gene mutation.[68,69]

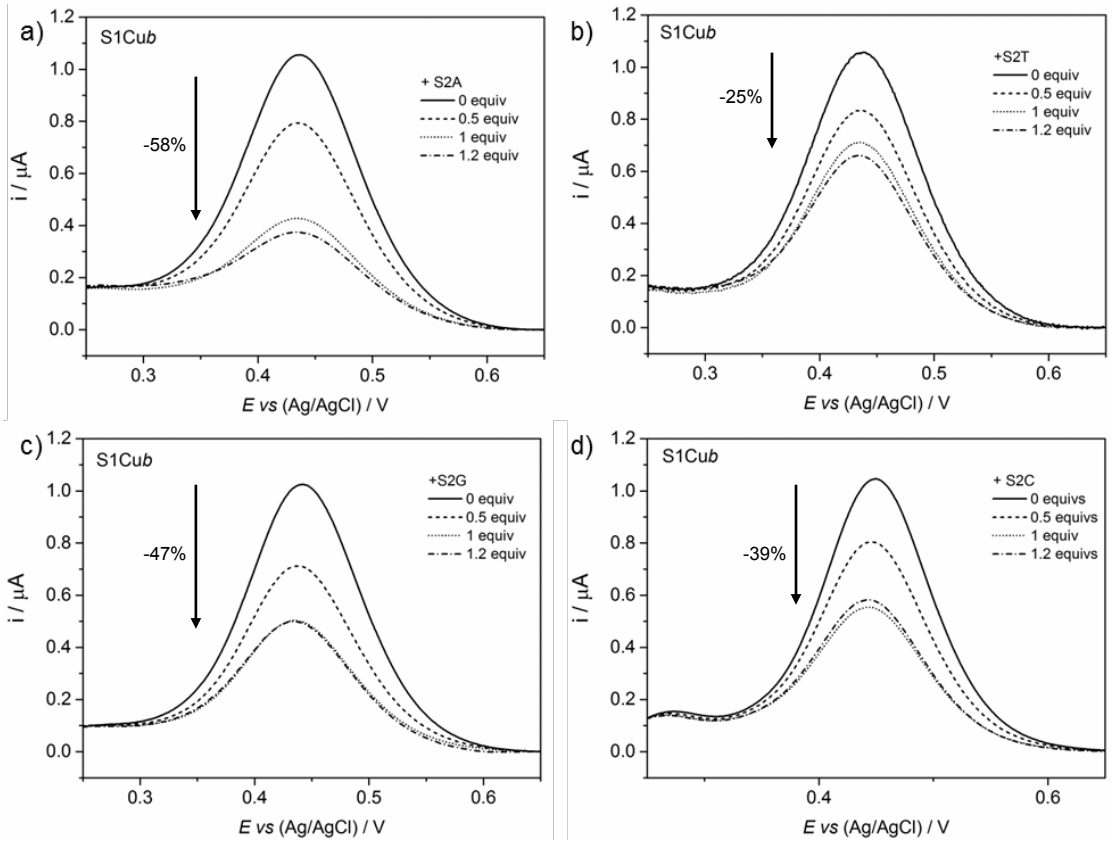


*Figure 5.26 - Square wave voltammetry examples of the observed decrease in currents for **S1Cub** when increasing amounts of **S2Y** are added. a) **S2A** b) **S2T** c) **S2G** d) **S2C**. 50 uM DNA, 100 mM NaCl, 10 mM Tris-HCl buffer pH 7.0. Data background subtracted and analysed using Origin Lab.*

*Table 5.9 - SWV sensing behaviour for **S1Cua** when hybridised with S2Y (X = T, A, G, C), and the Δ(ΔE$_p$) observed by CV for the hybridised strands.*

Target Strand	% Change in Current*	Δ(ΔE$_p$)* / *mV*
S2T	-25 (± 2.7)	7.0 (± 1.1)
S2A	-58 (± 0.6)	26.0 (± 1.5)
S2G	-47 (± 1.3)	9.7 (± 1.0)
S2C	-39 (± 2.5)	8.5 (± 1.7)

* Δ(ΔE*p*) = ΔE*p (single strand)* – ΔE*p (duplex). Values obtained at a scan rate of 15 mV s^{-1}. Change in current is measured at the E1/2* potential in the SWV. (.) shows the standard error of the mean of a minimum of three experimental repeats.

The question remains as to why the macrocycle displays differences in electrochemical output as a function of the identity of the nucleobase opposing the macrocycle in the target strand. Clearly size is not the only factor given that the size of T ≈ C and A ≈ G. With the observed current suppressions and the nature of intercalation, it is likely that electron transfer for the macrocycle is impeded by its embedding into the duplex core, with the degree of this intercalation dependent on the base opposite. Further evidence for this rationale is provided by the Δ(ΔE_p), the difference in peak separation between the duplexes **S1Cu*b*•S2Y** and the single strand. Similar trends to that observed for the current suppression were observed for Δ(ΔE_p) as shown in Figure 5.27. Peak separation as measured by CV, is a measure of the reversibility of a redox-active species. With the Δ(ΔE_p) changing with the opposing nucleobase it suggests that the base opposite imparts differing structural and or electronic properties onto the macrocycle and duplex. The previously discussed schematic (Figure 5.23) for the intercalation behaviour can perhaps be evolved. With the T_m data suggesting that the macrocycle stabilises the duplex more and also the electrochemical output is impeded less with pyrimidines opposite, it is possible that in order for the macrocycle to intercalate when A/G are opposite, it must cause a buckle in the duplex or kick-out the opposing base. In the case of both these compensations, the observed T_m reductions would be expected.

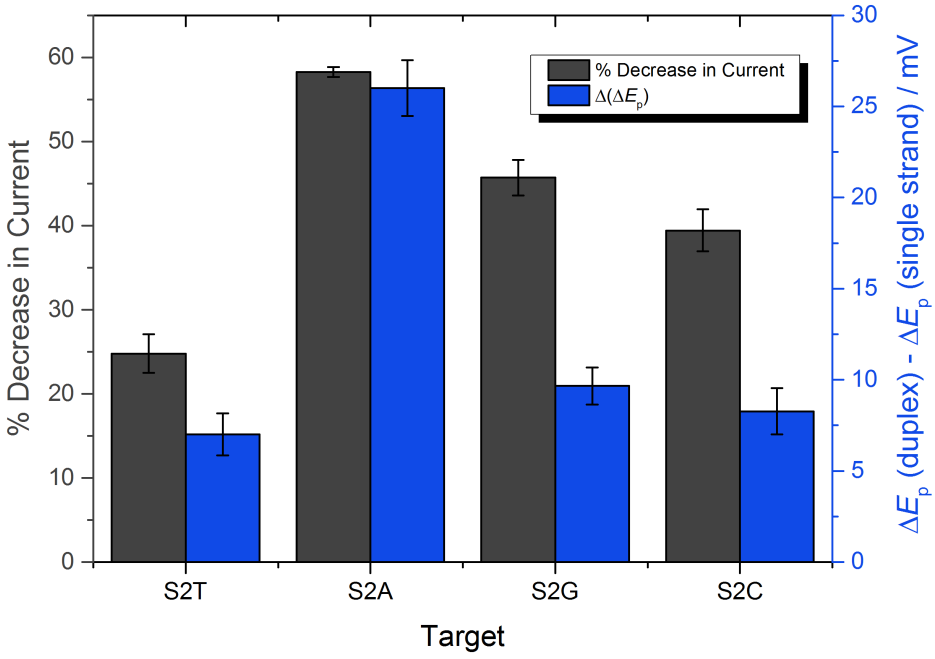


Figure 5.27 - The percentage current decreases (grey columns) and change in ΔE_p (blue columns) *when* ***S1Cub*** *is duplexed with* ***S2Y*** *(where X = T, A, G or C). 50 uM S1Cub, 1.2 equiv excess* ***S2Y****, 100 mM NaCl, 10 mM Tris-HCl pH 7.0. Error bars represent the standard error of the mean of 3 repeats.* $\Delta(\Delta E_p)$ *is measured at a scan rate of 15 mV* s^{-1}.

This rationale is supported by the molecular models shown in Figure 5.21, which suggests that a nucleobase might be 'flipped out' to accommodate the intercalating macrocycle **Cu*b***. Hence the degree of 'flipping out' is controlled by the size of the opposing nucleobase. Whilst this doesn't provide full rationale for these changes given the size similarities of A-G and T-C, this explanation supports both the melting and electrochemical data. A schematic simplified representation of this is shown in Figure 5.28.

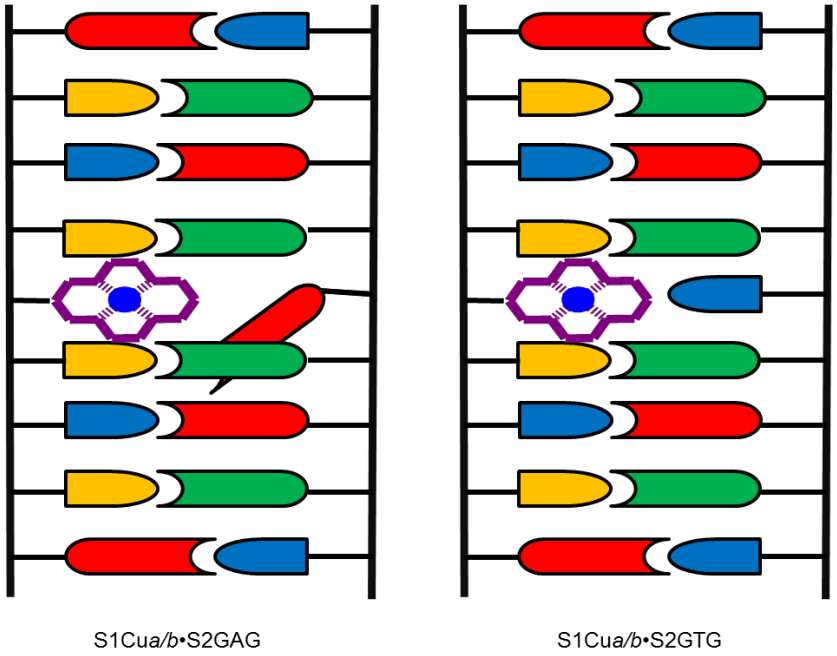


Figure 5.28 - Schematic representation of the 'flipping out' of a nucleobase to accommodate the intercalation of ***Cub.***

These systems were further studied by CD spectroscopy in order to establish any trends that might point to these findings. Figure 5.29(c+d) shows how the characteristic CD bands of the B-DNA and ICD of the macrocycle vary as the base opposite is changed. While no direct trend (T>C>G>A) is observed for changes in the ICD bands at 290, 317 and 358 nm, there are clear differences in the degree of induction of the respective transitions as a function of the nucleobase. For **S2C,** the largest differences are observed, with both an apparent increase in helical twist evidenced by larger B-DNA bands, but also with a larger ICD band at 290 nm. Interestingly, the other macrocycle transitions at lower energies appear similar to that of the **S2A** and **S2G**. This perhaps indicates that the specific electronic transition is aligned better with respect to the helical twist of DNA. A study of the duplex **S1Cu*b*•S2T** shows similar B-DNA intensities as those of **S2G** and **S2A** but the ICD bands at 317 and 358 nm are considerably enhanced. It is however apparent that for these lower energy transitions, the smaller pyrimidines impose an enhanced ICD, suggesting perhaps

that, as with the predictions from T_m experiments, the macrocycle can intercalate into the cavity more deeply. Although CD is not directly related to the degree of intercalation and is dependent on the angles of each transition with respect to the helical twist, and therefore cannot be used as a direct comparison with the electrochemical and T_m data to support structural rationales. It does however provide some insight into the differing abilities of the two stereoisomers to discriminate electrochemically the opposing nucleobase. The fast isomer (Figure 5.29 a+b), which is seemingly not able to discriminate between purine and pyrimidines, shows much smaller differences in these ICD bands, which is in line with observed electrochemical and thermal melting studies.

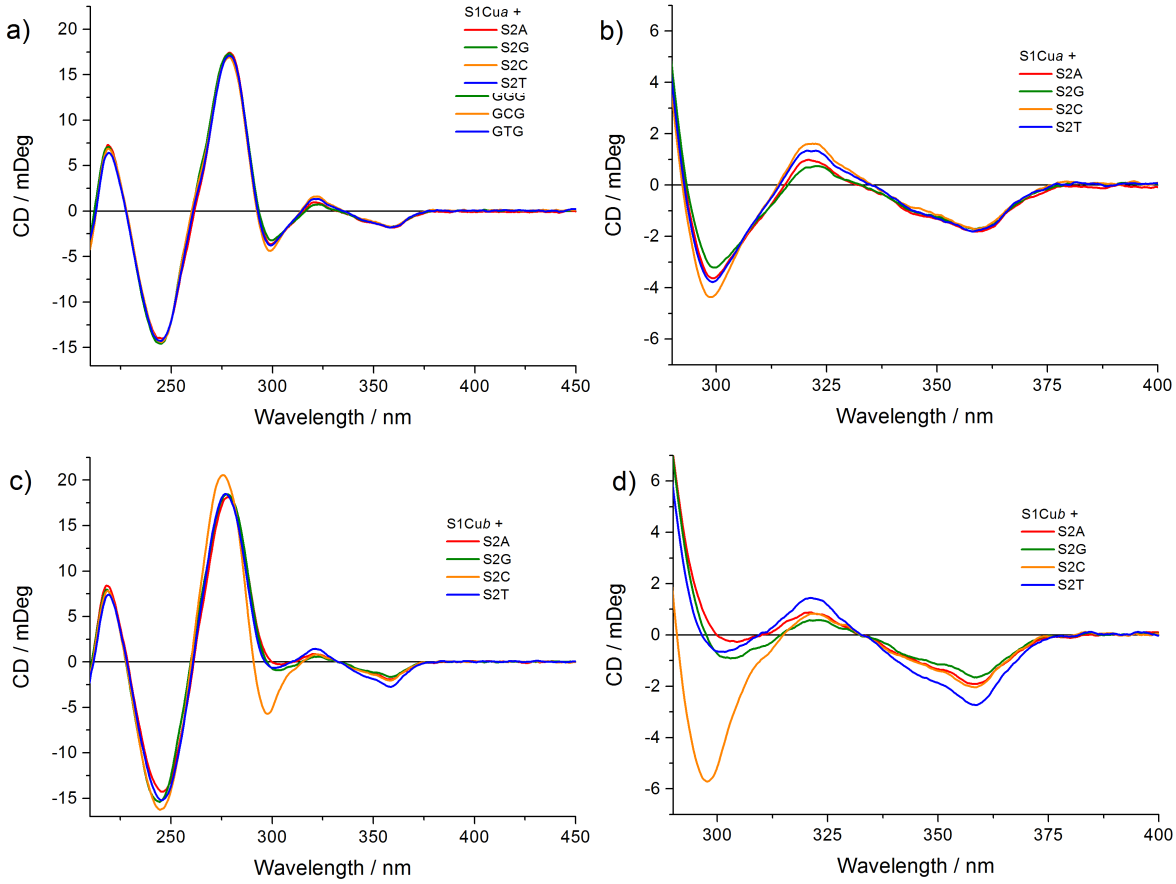


*Figure 5.29 - CD spectra of a) **S1Cua** and c) **S1Cub** when duplexed with **S2** with different canonical bases opposite. Shown also is a zoomed in region (290-400 nm) of the spectra showing differences in Soret and Q bands and helical structure for b) **S1Cua** and d) **S1Cub**. 5 μM of **S1** in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl.*

5.3 Conclusions

Nickel(II) and copper(II) tetrazacyclene macrocycles containing functionalities allowing for incorporation into DNA have been synthesised and used to tag DNA *via* automated DNA synthesis. This novel class of DNA modification represents the first successful incorporations of a planar macrocycle into the backbone of DNA, where the metal remains chelated to the complex throughout standard organic and DNA synthesis. Both linker/bridging and intercalators motifs show interesting thermal melting properties when replacing nucleobases within a 15-mer oligonucleotide, with the intercalator motifs as predicted inserting into the hydrophobic core of DNA. The electrochemical output of these complexes provides scope for future applications in the building and design of nanowires and redox sensors. In particular the use of an isolated stereoisomer of the intercalating motif S1Cu*a/b* allows for the discrimination of opposing nucleobases within a duplex *via* differences in electrochemical readout. The advantages of using electrochemically active redox probes for investigation of genetic material make this new class of DNA modification an attractive target for future studies. While the size of the opposing nucleobase is believed to be the main factor in the observed behaviour, it is not clear whether this is the only contributing property of such sensing. Other factors such as electronic properties, potential hydrogen bonding motifs and stacking contributions cannot as yet be ruled out. The base discrimination behaviour of the intercalating system represents an unprecedented example of this type of metal complex incorporated into the backbone of DNA, being capable of directly sensing minute changes in its immediate environment. This system has great potential for future use in SNP sensing assays.

5.4 Future Work

For the greater understanding of the sensing behaviour of these motifs I recommend that the following investigations could be carried out:

- Further CD analysis of intercalating motifs where the base opposite is removed, changed and studied using thermal ramps to investigate the binding modes of the intercalators. Molecular dynamic simulations that support these experiments would also be extremely useful.
- Expanding the SNP sensing profile of Cu*b* by investigating the ability to detect epigenetic changes in DNA sequences. Methyl-cytosine, hydroxymethyl-cytosine and oxo-guanine would represent interesting targets from a structural and diagnostic point-of-view, where these epigenetic changes have been associated with certain diseases.[57]
- Append intercalating motifs to an electrode surface via 5' or 3' DNA modifications for surface based sensing. Surface immobilised electrochemistry is shown to enhance sensing ability in reducing background interference, dramatically improving sensitivity and detection limits and requires far less DNA material. This step is vital for the future development of cheap, rapid and reliable SNP sensors for use in clinical or point-of-call devices.

Additionally 'zipper-like' stacking motifs involving multiple incorporations of metal complexes in to the DNA duplexes should be investigated further for possible conductance and hence for applications in the creation of DNA nanowires.

5.5 References

1. S. O. Kelley, N. M. Jackson, M. G. Hill and J. K. Barton, *Angew. Chem. Int. Ed.*, 1999, **38**, 941-945.
2. N. K. Modukuru, K. J. Snow, B. S. Perrin, A. Bhambhani, M. Duff and C. V. Kumar, *J. Photochem. Photobiol., A*, 2006, **177**, 43-54.
3. H. C. Becker and B. Nordén, *J. Am. Chem. Soc.*, 1999, **121**, 11947-11952.
4. Y.-H. Li, H. Long and F.-Q. Zhou, *Anal. Chim. Acta*, 2005, **554**, 86-91.
5. N. Thelwell, S. Millington, A. Solinas, J. Booth and T. Brown, *Nucleic Acids Res.*, 2000, **28**, 3752-3761.
6. M. J. Hannon, *Chem. Soc. Rev.*, 2007, **36**, 280-295.
7. Y. Saito, K. Motegi, S. S. Bag and I. Saito, *Bioorg. Med. Chem.*, 2008, **16**, 107-113.
8. N. Moran, D. M. Bassani, J. P. Desvergne, S. Keiper, P. A. S. Lowden, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2006, 5003-5005.
9. A. Okamoto, K. Kanatani and I. Saito, *J. Am. Chem. Soc.*, 2004, **126**, 4820-4827.
10. D. Boger, *Biorg. Med. Chem. Lett.*, 2001, **9**, 2511-2518.
11. O. Köhler, D. V. Jarikote and O. Seitz, *ChemBioChem*, 2004, **6**, 69-77.
12. Z.-Y. Zhao, M. San, J.-L. H. A. Duprey, J. R. Arrand, J. S. Vyle and J. H. R. Tucker, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 129-132.
13. J.-L. H. A. Duprey, D. M. Bassani, E. I. Hyde, C. Ludwig, A. Rodger, J. S. Vyle, J. Wilkie, Z.-y. Zhao and J. H. R. Tucker, *Supramol. Chem.*, 2011, **23**, 273-277.
14. S. Kumamoto, M. Watanabe, N. Kawakami, M. Nakamura and K. Yamana, *Bioconjugate Chem.*, 2008, **19**, 65-69.
15. J. Balintová, R. Pohl, P. Horákova, P. Vidláková, L. Havran, M. Fojta and M. Hocek, *Chem. Eur. J.*, 2011, **17**, 14063-14073.
16. B. Nordén and F. Tjerneld, *Biopolymers*, 1982, **21**, 1713-1734.
17. S. O. Kelley, J. K. Barton, N. M. Jackson and M. G. Hill, *Bioconjugate Chem.*, 1997, **8**, 31-37.
18. C. G. Pheeney and J. K. Barton, *Langmuir*, 2012, **28**, 7063-7070.
19. D. Kang, X. Zuo, R. Yang, F. Xia, K. W. Plaxco and R. J. White, *Anal. Chem.*, 2009, **81**, 9109-9113.
20. A. A. Lubin, B. Vander Stoep Hunt, R. J. White and K. W. Plaxco, *Anal. Chem.*, 2009, **81**, 2150-2158.
21. M. G. Hill, J. K. Barton, E. M. Boon, D. M. Ceres and T. G. Drummond, *Nat. Biotechnol.*, 2000, **18**, 1096-1100.

22. K. Hsieh, R. J. White, B. S. Ferguson, K. W. Plaxco, Y. Xiao and H. T. Soh, *Angew. Chem. Int. Ed.*, 2011, **50**, 11176-11180.
23. E. Farjami, L. Clima, K. Gothelf and E. E. Ferapontova, *Anal. Chem.*, 2011, **83**, 1594-1602.
24. Y. Xiao, X. Lou, T. Uzawa, K. J. I. Plakos, K. W. Plaxco and H. T. Soh, *J. Am. Chem. Soc.*, 2009, **131**, 15311-15316.
25. A. K. Boal and J. K. Barton, *Bioconjugate Chem.*, 2005, **16**, 312-321.
26. J. S. Swensen, Y. Xiao, B. S. Ferguson, A. A. Lubin, R. Y. Lai, A. J. Heeger, K. W. Plaxco and H. T. Soh, *J. Am. Chem. Soc.*, 2009, **131**, 4262-4266.
27. Y. Xiao, A. A. Lubin, B. R. Baker, K. W. Plaxco and A. J. Heeger, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 16677-16680.
28. Y. Xiao, A. A. Lubin, A. J. Heeger and K. W. Plaxco, *Angew. Chem. Int. Ed.*, 2005, **44**, 5456-5459.
29. *US Pat.*, US20150011712, 2013.
30. G. De Crozals, C. Farre, M. Sigaud, P. Fortgang, C. Sanglar and C. Chaix, *Chem. Commun.*, 2015, **51**, 4458-4461.
31. R. J. White, N. Phares, A. A. Lubin, Y. Xiao and K. W. Plaxco, *Langmuir*, 2008, **24**, 10513-10518.
32. A. Vallee-Belisle and K. W. Plaxco, *Curr. Opin. Struct. Biol.*, 2010, **20**, 518-526.
33. K. W. Plaxco and H. T. Soh, *Trends Biotechnol.*, 2011, **29**, 1-5.
34. S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam and R. Vilar, *Angew. Chem. Int. Ed.*, 2010, **49**, 4020-4034.
35. K. W. Jennette, S. J. Lippard, G. A. Vassiliades and W. R. Bauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1974, **71**, 3839-3843.
36. K. W. Jennette, J. T. Gill, J. A. Sadownick and S. J. Lippard, *J. Am. Chem. Soc.*, 1976, **98**, 6159-6168.
37. K. Suntharalingam, A. J. P. White and R. Vilar, *Inorg. Chem.*, 2009, **48**, 9427-9435.
38. K. Suntharalingam, A. J. P. White and R. Vilar, *Inorg. Chem.*, 2010, **49**, 8371-8380.
39. V. S. Stafford, K. Suntharalingam, A. Shivalingam, A. J. P. White, D. J. Mann and R. Vilar, *Dalton Trans.*, 2015, **44**, 3686-3700.
40. K. E. Erkkila, D. T. Odom and J. K. Barton, *Chem. Rev.*, 1999, **99**, 2777-2795.
41. I. Mames, A. Rodger and J. Kowalski, *Eur. J. Inorg. Chem.*, 2015, **2015**, 630-639.
42. V. W. W. Yam and K. K. W. Lo, *Coord. Chem. Rev.*, 1999, **184**, 157-240.
43. Y. Takezawa and M. Shionoya, *Acc. Chem. Res.*, 2012, **45**, 2066-2076.
44. J. Duprey and J. Tucker, *Chem. Lett.*, 2014, **43**, 157-163.

45. R. C. Holmberg, M. T. Tierney, P. A. Ropp, E. E. Berg, M. W. Grinstaff and H. H. Thorp, *Inorg. Chem.*, 2003, **42**, 6379-6387.
46. D. J. Hurley and Y. Tor, *J. Am. Chem. Soc.*, 1998, **120**, 2194-2195.
47. F. D. Lewis, S. A. Helvoigt and R. L. Letsinger, *Chem. Commun.*, 1999, 327-328.
48. S. I. Khan, A. E. Beilstein and M. W. Grinstaff, *Inorg. Chem.*, 1999, **38**, 418-419.
49. I. Vargas-Baca, D. Mitra, H. J. Zulyniak, J. Banerjee and H. F. Sleiman, *Angew. Chem. Int. Ed.*, 2001, **40**, 4629-4632.
50. K. M. Stewart, J. Rojo and L. W. McLaughlin, *Angew. Chem. Int. Ed.*, 2004, **43**, 5808-5811.
51. K. M. Stewart and L. W. McLaughlin, *J. Am. Chem. Soc.*, 2004, **126**, 2050-2057.
52. D. Magda, S. Crofts, A. Lin, D. Miles, M. Wright and J. L. Sessler, *J. Am. Chem. Soc.*, 1997, **119**, 2293-2294.
53. I. Grabowska, D. G. Singleton, A. Stachyra, A. Góra-Sochacka, A. Sirko, W. Zagórski-Ostoja, H. Radecka, E. Stulz and J. Radecki, *Chem. Commun.*, 2014, **50**, 4196-4199.
54. A. W. I. Stephenson, A. C. Partridge and V. V. Filichev, *Chem. Eur. J.*, 2011, **17**, 6227-6238.
55. D. Magda, M. Wright, S. Crofts, A. Lin and J. L. Sessler, *J. Am. Chem. Soc.*, 1997, **119**, 6947-6948.
56. N. Moran, D. M. Bassani, J.-P. Desvergne, S. Keiper, P. A. S. Lowden, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2006, 5003-5005.
57. J.-L. H. A. Duprey, Z.-y. Zhao, D. M. Bassani, J. Manchester, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2011, **47**, 6629.
58. B. Korybut-Daszkiewicz, R. Bilewicz and K. Woźniak, *Coord. Chem. Rev.*, 2010, **254**, 1637-1660.
59. A. Rybka, R. Koliński, J. Kowalski, R. Szmigielski, S. Domagała, K. Woźniak, A. Więckowska, R. Bilewicz and B. Korybut-Daszkiewicz, *Eur. J. Inorg. Chem.*, 2007, **2007**, 172-185.
60. I. Mames, U. E. Wawrzyniak, M. Woźny, R. Bilewicz and B. Korybut-Daszkiewicz, *Dalton Trans.*, 2013, **42**, 2382-2391.
61. W. Grochala, A. Jagielska, K. Woźniak, A. Więckowska, R. Bilewicz, B. Korybut-Daszkiewicz, J. Bukowska and L. Piela, *J. Phys. Org. Chem.*, 2001, **14**, 63-73.
62. H. V. Nguyen, Z.-y. Zhao, A. Sallustrau, S. L. Horswell, L. Male, A. Mulas and J. H. R. Tucker, *Chem. Commun.*, 2012, **48**, 12165-12167.

63. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, Wiley, 2000.
64. K. E. Erkkila, D. T. Odom and J. K. Barton, *Chem. Rev.*, 1999, **99**, 2777-2796.
65. K. Nakamoto, M. Tsuboi and G. D. Strahan, *Drug-DNA Interactions: Structures and Spectra*, Wiley, 2008.
66. C. Brotschi, A. Häberli and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2001, **40**, 3012-3014.
67. C. Brotschi and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2003, **42**, 1655-1658.
68. H. Davies, G. R. Bignell, C. Cox, P. Stephens, S. Edkins, S. Clegg, J. Teague, H. Woffendin, M. J. Garnett, W. Bottomley, N. Davis, E. Dicks, R. Ewing, Y. Floyd, K. Gray, S. Hall, R. Hawes, J. Hughes, V. Kosmidou, A. Menzies, C. Mould, A. Parker, C. Stevens, S. Watt, S. Hooper, R. Wilson, H. Jayatilake, B. A. Gusterson, C. Cooper, J. Shipley, D. Hargrave, K. Pritchard-Jones, N. Maitland, G. Chenevix-Trench, G. J. Riggins, D. D. Bigner, G. Palmieri, A. Cossu, A. Flanagan, A. Nicholson, J. W. C. Ho, S. Y. Leung, S. T. Yuen, B. L. Weber, H. F. Seigler, T. L. Darrow, H. Paterson, R. Marais, C. J. Marshall, R. Wooster, M. R. Stratton and P. A. Futreal, *Nature*, 2002, **417**, 949-954.
69. H. Namba, M. Nakashima, T. Hayashi, N. Hayashida, S. Maeda, T. I. Rogounovitch, A. Ohtsuru, V. A. Saenko, T. Kanematsu and S. Yamashita, *J. Clin. Endocrinol. Metab.*, 2003, **88**, 4393-4397.

Chapter 6

DNA Labelled Virus Particles

The work in this chapter was completed in collaboration with the Dafforn Group, School of Biosciences at the University of Birmingham and the Defence Science and Technology Laboratory (DSTL), Porton Down, UK. Parts of the results shown within were completed with or by Dr. Raul Pacheco-Gomez, Haydn Little and Nadja Steinke. Where appropriate, these are detailed within the text. I thank R.P-G, H.L and N.S for their valuable contributions and discussions related to this work. This work has recently been published in *ACS Synthetic Biology.*[1]

6.1 Introduction

6.1.1 Detection of Pathogens

The detection of microbial pathogens is an essential part of diagnosis and treatment in healthcare. Rapid detection and identification of bacterial strains exhibiting antibiotic resistance steers decisions on which class of antibiotic should be chosen for treatment. Microbial pathogen detection has also played an important part in the food industry where detecting crop diseases reduces raw material costs while monitoring hygiene during processing increases safety and reduces waste. Furthermore recent outbreaks of diseases such as Ebola and the growth in world travel have increased interest in the use of pathogen detection systems at border controls for national security.[2,3]

Most microbial pathogen tests rely on detecting three underlying aspects of microbial life: firstly the ability to grow and replicate; secondly the presence of microbial specific antigens and thirdly the presence of DNA from the pathogen. Detection of microbial pathogens by virtue of their ability to replicate has been a core part of pathogen detection for more than 100 years. Suspicious samples are cultured in a range of growth media, with the pathogen then identified by a combination of parameters including media selectivity, colony morphology and microscopic anatomy. Assays that detect the presence of antigens specific to the pathogen rely on antibody:antigen recognition. Such immunoassays include laboratory-based methods such as ELISA and close-to-use methods e.g. lateral flow devices. Methods for detecting regions of DNA that are specific to the pathogen are normally carried out in conjunction with the polymerase chain reaction (PCR) where genetic material is amplified using a thermo stable DNA polymerase. This method is particularly attractive where pathogens are present at very low levels. Despite the excellent sensitivity of this technique, transduction of information regarding the success of these amplifications is often time-consuming, laboursome and requires highly trained technicians operating in a controlled environment.[4] For example, gel electrophoresis, arguably the most popular method for this

analysis during the 1990's and still regularly used today, requires the use of high voltages combined with carcinogenic dyes to provide an answer.[5]

Much work has been carried out in recent years to somewhat overcome these problems. One of the major breakthroughs in this area being the development of real-time PCR (RT-PCR/ qPCR).[5-8] This technique uses a range of methods to detect the production of the amplicon of PCR in real time, usually based on dye labelled oligonucleotides. While this technique is highly accurate it requires expensive equipment and reagents and is prone to mistakes and false positives due to the extreme sensitivity to changes or contamination. Sample preparation is crucial to its success and usually requires highly trained technicians to carry out this technique and in turn interpret its complex data outputs. The sensitivity to contamination also means that analysis of samples in real time is not a reality, as pre-purification is essential to ensure the successful quantification. While RT-PCR has been ground-breaking in terms of the speed of genomic mapping and detection in the laboratory its drawbacks have meant its use in clinical assays and in-the-field devices has yet to be developed.

6.1.2 M13 Bacteriophage

6.1.2.1 Structure and Properties

M13 Bacteriophage (M13) (depicted in Figure 6.1) is a lysogenic virus particle that infects *E.coli.* It is a long fibrous rod-like nanoparticle (*ca*. 880 nm in length. 6.6 nm wide)[9] that houses a viral ssDNA that is 6407 nucleotides long. The ssDNA is contained within a self-assembled protein network of 5 coat proteins (p8, p6, p3, p9, p7). The shaft of the virus is made up of 2700 copies of a major coat protein p8/pVIII, a 50 amino acid alpha helical peptide. While the other capping proteins, which are only made up of 30-40 amino acids play roles in the viral function and interaction with *E.coli.*

M13 can be easily produced in milligram quantities by infection of *E.coli* and a few simple purification steps that make it ideally suited as a bio-scaffold. This factor on top of its uniformity and relative ease of modification through changes in coat protein genes has led to M13 bacteriophage along with other bacteriophages becoming highly popular in phage display technologies.

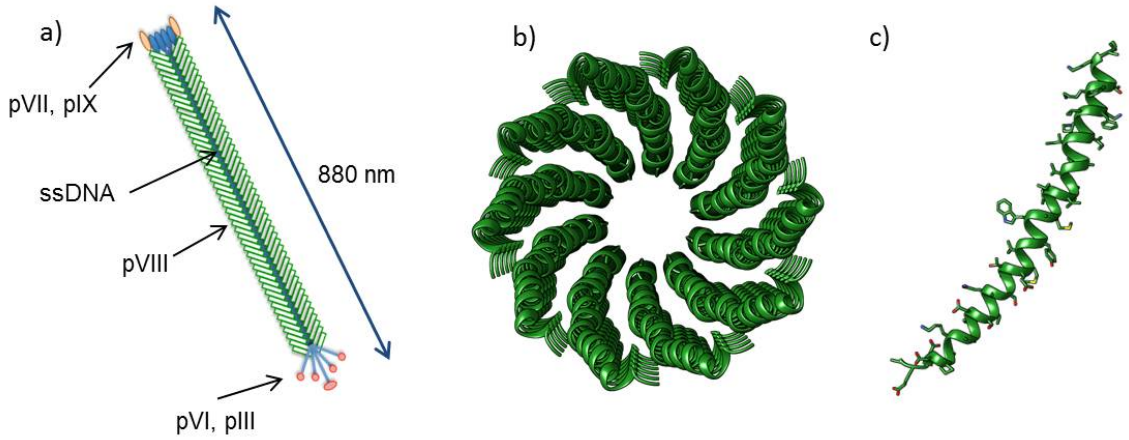


Figure 6.1 - a) schematic representation of a cross-section of M13 and its coat proteins. b) self-assembled pVIII sheath that encapsulates ssDNA. c) pVIII protein and side chains.

More interestingly to the chemist's toolbox are the pVIII protein structure and its position within a self-assembled tube of 2700 copies which bear solvent accessible side chains (See Figure 6.2) that can be functionalised using simple conjugation chemistries. The synthetic flexibility afforded by these proteins means M13 has in recent years been utilised as a nanoscaffold for a whole host of applications.

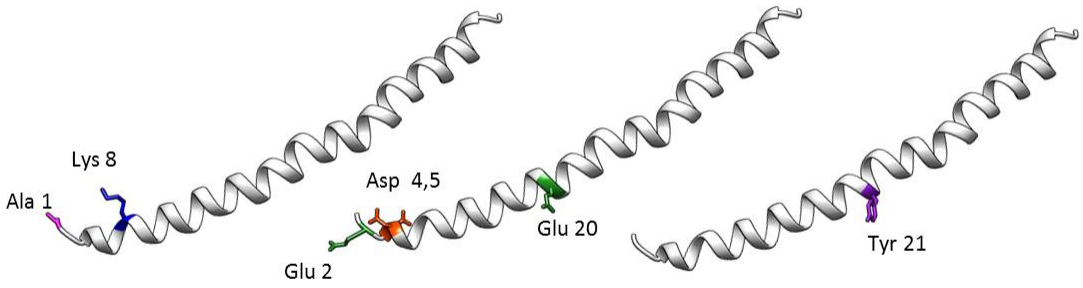


Figure 6.2 -The reactive side chains of the pVIII protein of M13 that allow for specific conjugation to its exterior.

6.1.2.2 Applications and Uses

Given the uniformity and availability of functional groups on this remarkable nanoparticle, it provides a rod-like nano-scaffold that can be functionalised to create new materials and devices. Many developments of these functional biomaterials have been driven by Angela Belcher (M.I.T) and co-workers, who have successfully established a range of novel materials for a whole host of applications. M13 scaffolds with genetically engineered coat proteins and M13 stabilised carbon nanotubes have been utilised in targeting and imaging of cancer cells.[10-12] Furthermore bio-templation and nucleation of transition metal oxides (manganese,[13] cobalt,[13,14] gold,[14] iridium[15]) using engineered pVIII peptides that bind specific metals have found interesting applications in the development of lithium batteries and energy storage. Additionally bio-templation approaches have been used to create magnetic single crystals and semiconducting nanowires by nucleation of ZnS, CdS, CoPt and FePt respectively.[16] Her group has also exploited the accessible nature of the lysine and N terminus of wtM13 by attaching porphyrin motifs to the major pVIII coat with a view to manipulating energy transfer processes in light harvesting systems that could mimic photosynthesis (Figure 6.3b)[17]

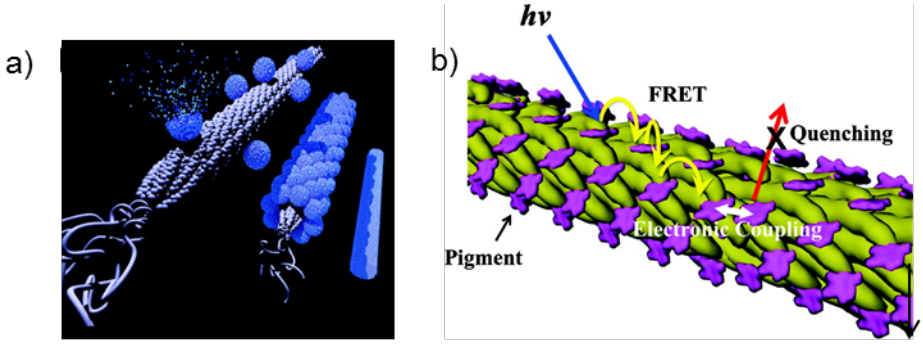


Figure 6.3 - Schematic representations of bio-templated nanowires[16] and b) M13 templated light harvesting nanoattenae,[17]

Furthermore Jennifer Cha (Boulder) *et al* have exploited M13 in diagnostics and sensing, which has seen the development of ELISA type assays[18] and protein detection motifs[19] that rely on the M13 rigidity and the ease of conjugation to its coat proteins.

6.1.3 Linear Dichroism

Polarised light spectroscopy is most commonly associated with circular dichroism (CD) spectroscopy which has been instrumental in understanding bio-molecule structure and in monitoring biomolecule-ligand interactions (Chapter 2.2.2).

Linear dichroism (LD) spectroscopy however, differs from circular dichroism spectroscopy in that it utilises linearly polarised light (as opposed to circular) to monitor structure. Linearly polarized light is light that is polarized along a single plane of directionality. Polarising light in this manner produces an electric field vector/wave, which oscillates in a given direction. By monitoring differences in the absorbance of parallel and perpendicular polarized light to the orientation of a light absorbing species/molecule, it is then possible to deduce information regarding the arrangement of molecules.[20]

When a molecule absorbs light, an electronic transition occurs in which electron density within the molecule is excited and shifted. The direction of these shifts is known as the

electronic transition polarisation. This is the property which LD is able to measure and provides insight into structure and position of absorbing species either in complex biomolecules or in simple molecules. Important to the success of this technique is that molecules of interest display electronic transitions that are orientated in some way to the directionality of the parallel or perpendicular light. Where the LD signal can be written as;

Equation 6.1[20]

$$LD = A_{||} - A_{\perp}$$

$A_{||}$ is the absorbance measured parallel to the orientation axis and $A_{\perp}$ is the absorbance perpendicular to that axis.

In a solution, molecules are orientated randomly and therefore the absorbances have no directionality and either no LD or a very small LD signal is observed depending on the specific interactions between molecules. The electronic transitions are essentially cancelling each other out. It is therefore important to align molecules of interest in order to produce LD signals from which structural information can be deduced. One of the major advantages of LD over other spectroscopic techniques (UV-vis and CD) is that only aligned/orientated samples will produce a signal; which reduces the background noise or false positives that could be associated with contaminants. This trait means LD has great scope for use in on-site rapid use assays where cumbersome purifications steps are not possible.

6.1.3.1 Alignment methods

There are many methodologies that can be used to align a molecule so that they produce an LD signal. The detail of all the available methodologies is beyond the scope of this work, but includes the use of magnetic and electric fields, gel, liquid crystal and evaporative techniques.[20] Described below are the two most popular approaches utilised for the orientation of biomolecules, stretched film and flow alignment. The latter providing the basis for this project.

Stretched film

Stretched film orientation of molecules utilises a polymer film (usually polyethylene or poly(vinyl)alcohol) to orientate sample molecules. By placing the sample on top of the film and stretching the film in a singular directionality, the sample molecules are orientated in the direction of the stretch of the polymer. This then allows LD signals to be observed. Stretched film has been popular for samples where shear flow alignment (*vide infra*) is not possible, it is used generally for, but not limited to, small organic molecules such as organic dyes.[21,22] Nucleic acids[23] and bacterial protein reaction centres have also been studied.[24,25]

Flow Alignment

If a sample is dissolved in solution (the ideal for understanding *in vivo* function and structure of biomolecules) and the solvent flowed at a sufficient rate past a static surface, then molecules can experience shear forces large enough to orientate them along the directional axis of the imposed flow.[20,26]

By placing a sample between a spinning quartz tube and a stationary quartz rod, the molecules of interest can be aligned in the direction of the flow if the morphology or structure of the sample allows for uni-directional orientation in flow (Couette flow). This property is usually reserved for fibrous like biomolecules such as bacteriophages and large strands of DNA/RNA. Couette flow is currently the only method for studying the LD of samples in solution, but is particularly advantageous when it can be used due to requiring only very small samples volumes (20-60 µL). Additionally, by spinning a sample in a cylindrical vessel the sample constantly recirculates and therefore multiple accumulations can be taken to reduce the signal-to-noise ratio. The first examples of the use of Couette flow to study nucleic acid-intercalator interactions was described in 1976 by Nordèn and Tjerneld; they showed how calf-thymus DNA (ctDNA) could be aligned using this method.[27] In the years since, Couette flow has been used to study a whole wealth of biomolecules using LD in order to better understand their structure. Accordingly flow-aligned LD has been performed on a

range of alignable structures including nucleic acids,[23,27,28] protein fibres,[29-33] carbon nanotubes[34,35] and membrane proteins.[36-38]

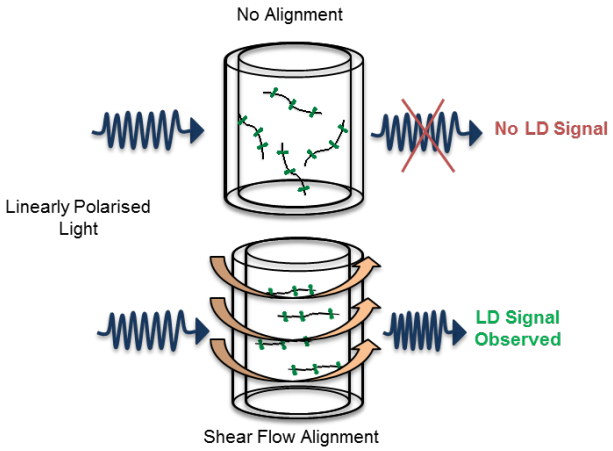


Figure 6.4 - Schematic representation of flow aligned LD spectroscopy using Couette flow.

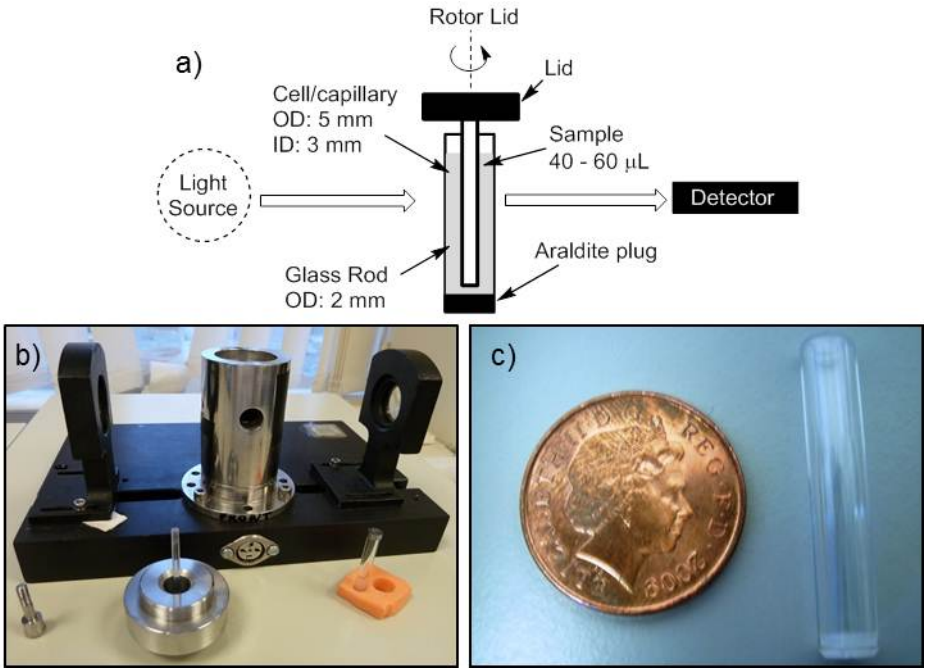


Figure 6.5 - Images of the Couette cell apparatus used for flow alignment LD. a) Schematic representation of the Couette cell, b) photograph of the cell holder and c) a size comparison of the Couette cell.

6.1.3.2 Linear Dichroism and M13 Bacteriophage

Given the fibrous properties of M13 bacteriophage, it was rationalised as a potential candidate for studying using LD. Clack *et al* showed in 1992 that bacteriophages like M13

could be aligned in extension pulse flow and hence produced a strong LD signal as the chromophores align in a unidirectional manner.[39] As previously outlined, if the absorbances and hence electronic transitions occur with some parallel or perpendicular character with respect to the orientation axis, then an LD signal is observed (see Figure 6.4). Rod-like particles such as M13 are known to tumble in solution in a Brownian motion. The LD signal observed in Couette flow is the result of biasing the distribution of particle alignment in the direction of flow, the level of bias can be quantified by the rotational Peclet number.[40]

In 2012, modified M13 bacteriophages were shown to align by Dafforn *et al* using considerably lower shear velocities than previously used,[41] the signals were however smaller (LD is proportional to the alignment) but significantly larger than the background. The sample volume required was significantly reduced (50 µl *vs.* 30 ml) meaning potential assays/devices could work on a more realistic scale for biomolecules. From the observed LD signal (Figure 6.6), a number of conclusions regarding M13's structure could be made.

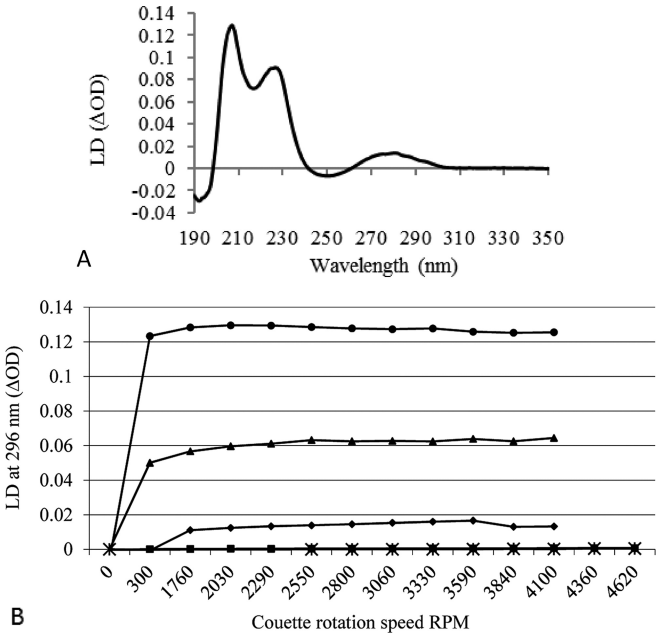


Figure 6.6 - (A) LD spectrum of wild-type M13 (wtM13) orientated using Couette flow at 3000 rpm (0.7 µM) (B) Relationship between concentration of wtM13 and rotation speed on observed LD signal at 296 nm. Circles - 0.56 µM, triangles - 0.28 µM, diamonds - 0.07 µM, squares – 5.6 nm.[41]

The spectrum consists of a number of regions of interest in both the near-UV (260-350 nm) and far-UV (190-260 nm) attributed to aromatic transitions of tryptophan and tyrosine residues and electronic transitions of the peptide backbone respectively.[39,41] Full details of these transitions will be discussed in Section 6.2.2.3.

Owing to the alignable nature of wtM13, Pacheco-Gomez *et al* utilised this bio-scaffold to design an immunoassay based on antibody:antigen recognition. By genetic modification of the minor coat protein pIII to produce a fused antibody recognition site for *E.coli* O187, the group were able to sense pathogenic bacterium in concentrations of x 10^7 cells per ml.[41] Although the detection limit is considerably higher than would be required for laboratory based assays it showed great scope for using M13 as a bio-scaffold for a whole host of detection assays with views to creating multi-modality. Since the publication of this work, the group has been successful in reducing the detection limits 100-fold by changing the antibody fusion site to the more abundant pVIII location.

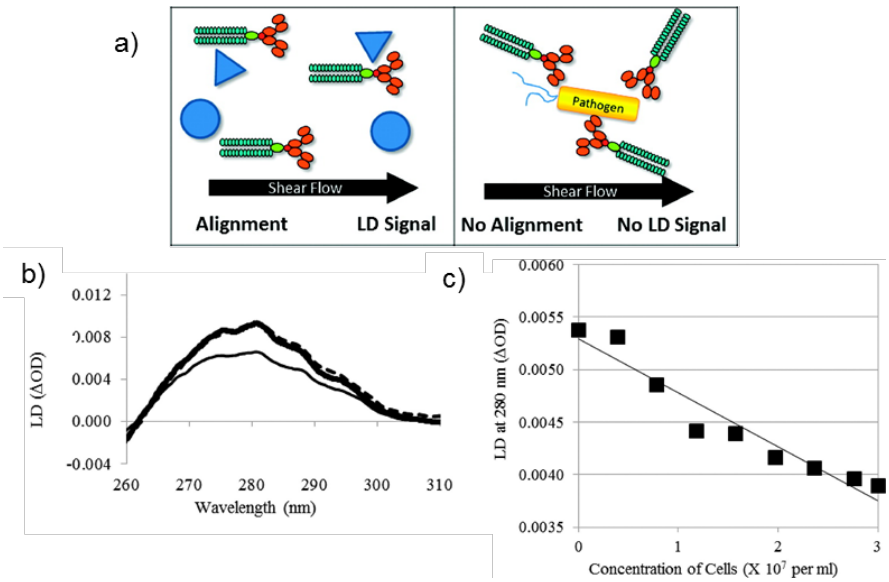


Figure 6.7 - a) Schematic representation of an antibody:antigen M13 based immunoassay. b) Change in LD spectrum at upon addition of E.coli O187 (before – thick solid, after addition – thin solid). c) Concentration dependence on LD sensing of E.coli O187 in x 10^7 cells per ml. Edited with permission from ref.[41].

6.1.3.3 Polymerase Chain Reaction

The invention of the polymerase chain reaction (PCR) has been so valuable to the biological sciences that it is hard to imagine how our understanding of molecular biology and biological processes would have developed without it. So important was the development of this technique that K. B. Mullis was awarded the Nobel Prize in Chemistry in 1993.[42] Although he was not the first to consider the repair principle of DNA replication, described by Khorna *et al* in 1971,[43] he is considered to be the inventor of the PCR technique. Although the description of which was first published in 1985,[44] the major breakthrough came in the following years. Mullis *et al* were credited with the improvement of what was considered to be an inefficient process through the purification and use of a thermostable DNA polymerase, *Taq* polymerase (isolated from *Thermus aquaticus*). The enzyme found in hot springs and thermal vents was discovered and found to be able to withstand high temperatures, that of which would usually denature a protein.[45]

The PCR process, shown schematically in Figure 6.8, is designed as such that it can exponentially increase/amplify a specific region of DNA through the use of short oligonucleotide primers that bind to a DNA target/template and catalyse the polymerisation and extension of DNA. A typical PCR reaction contains very few reagents; a DNA template, DNA polymerase, forward and reverse primers, deoxynucleoside triphosphates (dNTPs), and a buffer solution. PCR is considered to have three main steps used in cycles to achieve maximum amplification;

1. Denaturation (A) – in which elevated temperatures (generally > 90 °C) are used to disrupt hydrogen bonding in double stranded DNA templates.
2. Annealing (B) – temperature is lowered to 45-65 °C in order for the primers to bind to a single stranded target. The excess of these primers drives forward the binding over large complementary single strands that were previously denatured.

3. Elongation (C) – temperature is again fine-tuned, specifically for the polymerase used. This allows for the optimum efficiency of the enzyme. The enzyme, driven by the newly formed primer-template double strands, polymerises and extends the primer strand using the available dNTPs by condensation of the 5' phosphate to the 3' OH of the DNA.

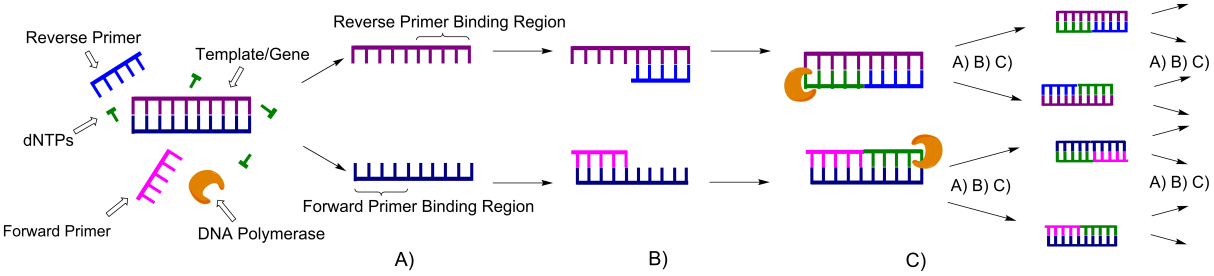


Figure 6.8 - Diagrammatic representation of a standard PCR reaction utilising forward and reverse primers. A) Denaturation. B) Annealing. C) Elongation

6.1.4 Conclusions and Project Aims

With current technologies available for the detection of genetic material indicative of a pathogen being mainly limited to either laboursome or highly sensitive PCR based methodologies, it was clear there was room for the development of new portable technologies that could yield the same information, but faster and with less reliability on the user input.

The desire to create faster, ever more reliable assays for detection of pathogens has led to the development of a new bio-scaffold based around M13.[41] In this project we aimed to utilise this scaffold again alongside LD, by conjugating short oligonucleotide targets to the surface of M13. These oligonucleotides would be specific to a pathogen of interest, and be capable of binding genetic material that could be found perhaps at a crime scene, in a blood sample of a patient or perhaps even on the battlefront.

The combination of M13, DNA recognition and LD would hopefully in turn generate a new sensor platform that can broaden the scope of M13 based multi-modal assays. The work shown herein, was completed as part of a collaboration with the Defence, Science and Technology Laboratory (DSTL) and shows the first proof-of-concept design of a M13-DNA conjugate capable of amplifying genetic templates through an unprecedented example of PCR on a virus, followed by rapid detection of the amplicon and hence presence of a pathogen by linear dichroism spectroscopy.

6.2 Results and Discussion

6.2.1 Design Strategy

Essential to the success of the assay development were four key factors;

- That oligonucleotides could be appended to the exterior of wtM13 either through direct protein-DNA conjugations or interactions.
- That M13 bacteriophage is able to withstand the extremities of PCR amplification.
- That changes to the hydrodynamics of M13-DNA scaffolds following a PCR amplification could be detected using LD, given the development of small portable LD devices currently in prototype phase.
- Detection is specific to the pathogen or genetic material of interest.

The design should also be as such that synthesis and purification of these systems is not particularly expensive or time-consuming and the results produced by such changes could be easily interpreted by non-experts or untrained personnel. Figure 6.9 below shows a diagrammatic representation of the assay design and step-wise approach to biosensing of pathogens. Where an M13-DNA conjugate can replace a forward or reverse primer in a standard templated PCR reaction (Figure 6.8) to produce an alignable scaffold that can detect amplification through changes in its orientation axis.

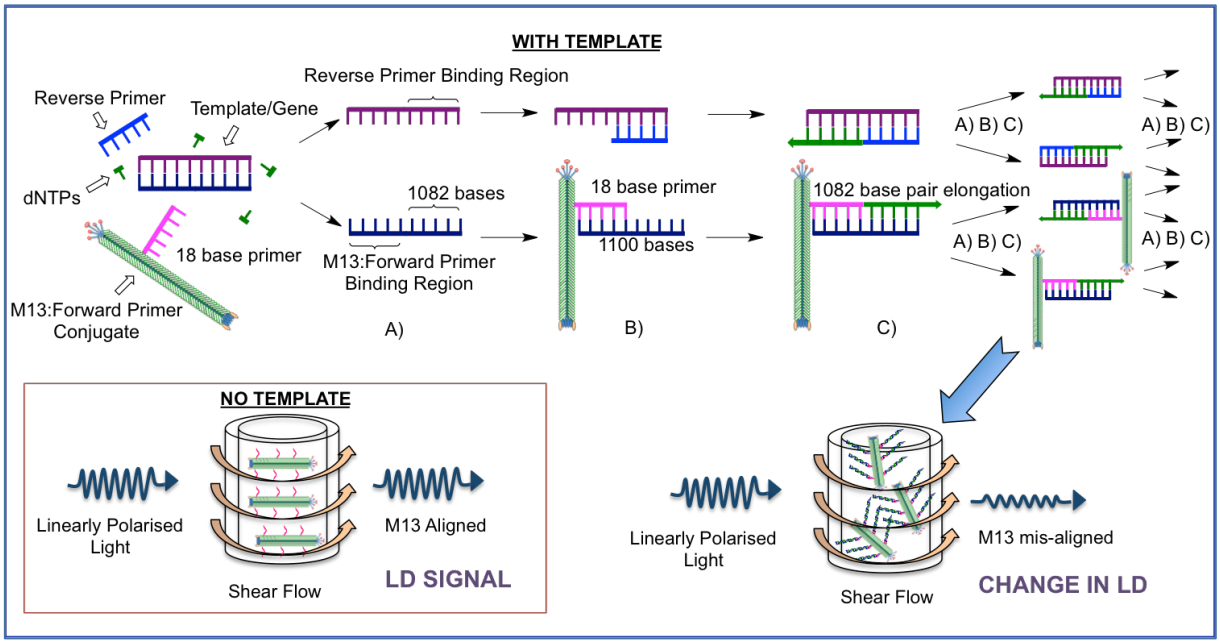


Figure 6.9 - Diagrammatic representation of the concept behind a M13-PCR scaffold in both amplification and detection of pathogenic DNA. A) Denaturation, B) annealing andC) elongation.

The ampicillin resistance gene (*ampR) of the Escherichia coli* plasmid *pBR322 w*as chosen as a target for the PCR assay.[46] Its role in coding for the production of β-lactamase, a protein responsible for catalysing the hydrolysis of penicillins, has been widely used as a selective gene marker in biotechnology. This, coupled with its banned usage in modified foods owing to fears over horizontal gene transfer to pathogens makes it an interesting target for this study.

The oligonucleotide primers were therefore designed to be specific to this gene/template according to the sequence published by Sutcliffe in 1978,[46] and were as follows:

- Forward primer – 5'-ATGAGTATTCAACATTTC-3'
- Reverse primer – 5'-TTACCAATGCTTAATCAG-3'

For purposes of this study the forward primer was used in conjunction with M13. It is important to the success of the design that the M13 surface can be modified with a primer

sequence, and that the structure can be characterised to determine conjugations and behaviour. We therefore chose to functionalise the oligonucleotides with a thiol moiety to allow for conjugation and TAMRA, a rhodamine derivative, as a chromophore and fluorophore to aid in the study of these systems. This forward primer (**R-For**) was synthesised using ULTRA-mild automated oligonucleotide synthesis (detailed in Chapter 7.2.2) using commercially available 5' thiol functional functionality and a thymidine substituted TAMRA derivative (see Figure 6.10). Additionally, to probe the binding ability and validity of the design two strands were synthesised, a short complementary (**F-Com**) and non-complementary (**F-Ran**) DNA targets functionalised with fluorescein (**6-FAM**) to allow for fluorescence spectroscopy analysis. Details of these strands are listed in Table 1 below. Full details of the characterisation of these oligonucleotide strands are available in appendix 19.

Table 1 - Synthesised oligonucleotides.

Strand	Sequence (5'-3')**
R-For	HS-C6-(TAMRA-dT)-ATGAGTATTCAACATTTC
F-Com	(6-FAM)-GAAATGTTGAATACTCAT
F-Ran	(6-FAM)-TCATCAGTCAGTCAGTCA

**Modifications and there structures are shown in Figure 6.10.

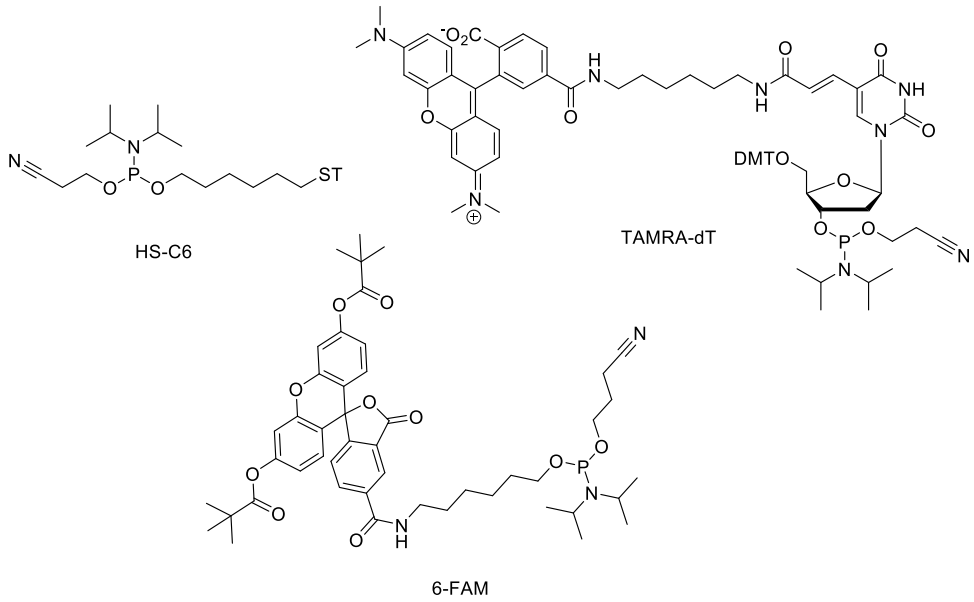


Figure 6.10 - Phosphoramadite monomers for oligonucleotide modifcation. T = trityl, DMT = dimethoxytrityl.

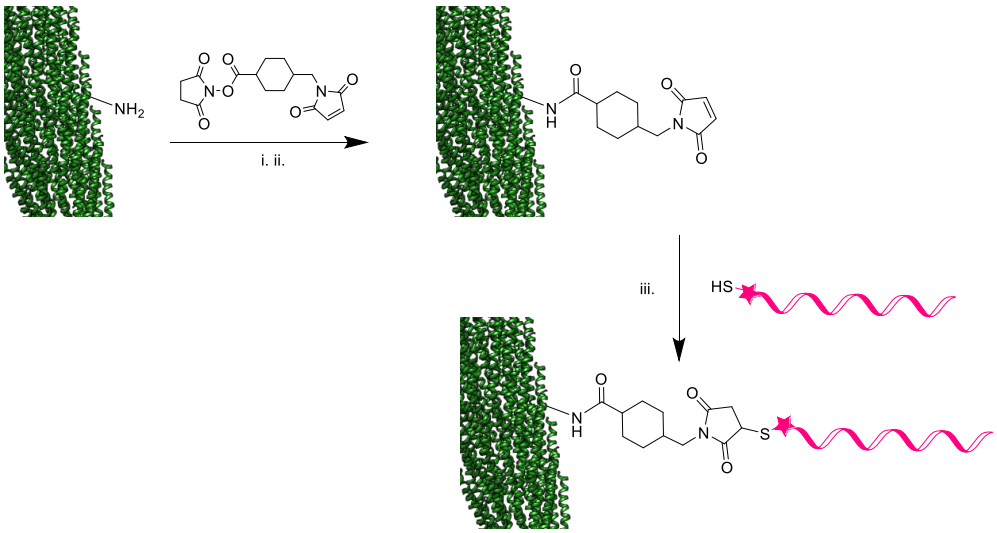
6.2.2 Conjugation, Purification and Characterisation

6.2.2.1 Conjugation

Research conducted previously by the Dafforn Group indicated an enhanced immunoassay sensitivity was achieved when protein pVIII was functionalised over pIII. This coupled with pVIII's greater abundance made it a suitable choice for conjugation.

An added advantage alongside the abundance of this protein was the accessibility of the functional groups as discussed in 6.1.2.1. The approach was via a heterobifunctional linkage that is commonly used in bioconjugation chemistry, and had previously been used by the group[47] in conjugating DNA to M13.[19] Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) is a water insoluble linker bearing an amine reactive NHS ester, a cyclohexyl spacer, and a reactive maleimide for conjugation to thiols. This linker forms stable amide bonds, leaving the reactive maleimide to react second in a two-step reaction. The N-terminal alanine residues and lysine side chains, of which there are approximately 2700 of each, were an ideal target.

With some optimisation, the conjugation was successful. Firstly, thiolated DNA was deprotected of its trityl protecting group using $AgNO_3$ and DTT and stored with a twofold excess of TCEP, a reducing agent specific for disulfides. This was to ensure the maximum amount of thiolated DNA was available for the reaction. Secondly, a 2 mg/ml solution of wtM13, which had been grown by Dr. Pacheco Gomez *via* a previously described procedure [41] was reacted with a 1000-fold excess (over pVIII protein – 2700 copies) for 1 hour. The lower pKa and enhanced solvent accessibility of the terminal amine results in a preferential conjugation over the lysine residue (Lys 8 - Figure 6.2).[19,48]



*Scheme 6.1 - Conjugation of **R-For** to the M13 major coat protein (pVIII) via a heterobifunctional linker, SMCC. i) SMCC in DMSO, 1 hr, RT. ii) glycine, 15 min, RT. iii) **R-For**, TCEP, 16 hrs, 5 °C.*

Unreacted NHS esters were then reacted with glycine for 15 minutes before passing through a sephadex de-salting column to remove unwanted organics and salts. SMCC labelled M13 was then reacted with the deprotected thiol DNA (**R-For**) overnight (Scheme 6.1). **R-For** was added in a 100-fold excess of the modified M13 with the aim to conjugate multiple strands to each bacteriophage particle.

6.2.2.2 Purification and Characterisation

The M13-DNA conjugates (M13-**R-For**) were purified and isolated from unreacted components using size exclusion chromatography (SEC). Figure 5.11a shows a typical SEC chromatogram of a DNA conjugated M13. The modification of the oligonucleotide with TAMRA allowed for the tracking and confirmation of conjugation via a distinctive absorption at 555 nm (pink). The three absorbances; 269 nm for a combined DNA-protein from the M13 conjugated DNA and internal ssDNA; 280 nm for the pVIII protein and 555 nm for TAMRA are all observed in the first two peaks eluted. The first peak collected was adjudged to be that of the conjugated material, because as the largest molecule it would be expected to be eluted because larger molecules should move faster through the porous sephadex column,

and secondly because of the ratio of the 269:280:555 nm peaks to each other (note in the 555 nm signal is multiplied by 10). The second peak is attributed to unreacted **R-For** and the third to small organic molecules such as TCEP that were used during the conjugation process. Additional confirmation of conjugation came from sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) shown in Figure 5.11b. Denaturation of the M13-**R-For** conjugate before loading allowed for the separation of the major coat protein pVIII. A characteristic pVIII band[49] observed in both wtM13 control and M13-**R-For** conjugates was observed, along with a second band (*ca*. 12 kDa), only in the case of conjugation and is highly indicative of a successful conjugation of **R-For** (6782 Da).

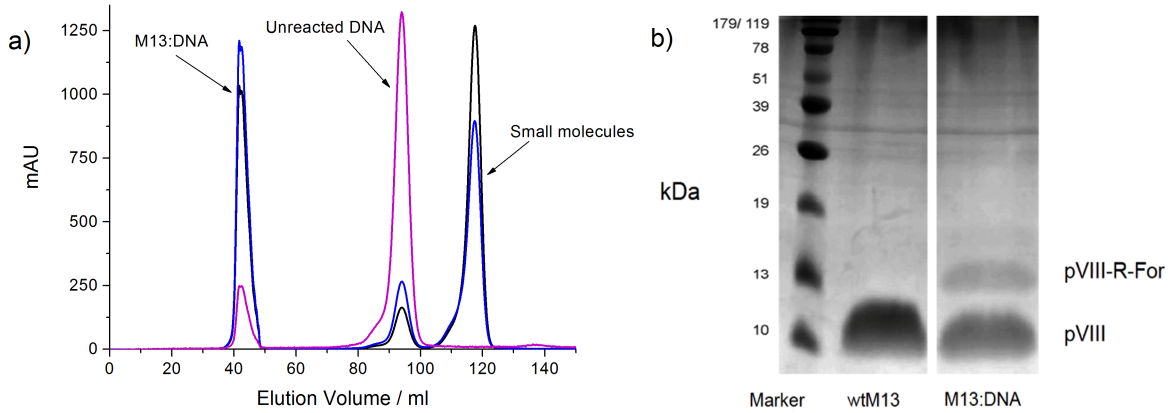


*Figure 6.11 - a) Size exclusion chromatography of **M13-R-For** conjugate monitored by three wavelengths. Blue - 269 nm, black - 280 nm and pink - 555 nm x 10. b)SDS PAGE analysis of the pVIII protein conjugated to **R-For** shown alongside wtM13. PAGE analysis carried out by Haydn Little.*

Measurement of the UV-Vis spectrum (an example is given in Figure 6.12) of the isolated M13-**R-For** conjugates provided further evidence for the successful conjugation, with a characteristic TAMRA absorbance observed at 556 nm (λ_{max}). The benefit of this fluorophore became apparent when it was possible to use this absorbance, which is uncomplicated by the DNA and peptide transitions, to quantify the conjugation efficiency.

From this quantification it was determined that 0.3% of the pVIII proteins were conjugated to **R-For**, which meant an average of around 8 oligonucleotides attached to each M13 particle (it was assumed that each M13 has exactly 2700 copies of pVIII). A low amount of oligonucleotide appended to each bacteriophage was not unexpected, as only an excess large enough to conjugate 4% of the bacteriophage coating was reacted. Low coverage was part of the design strategy as previous work within the group had shown this to be effective in ensuring an unperturbed LD signal for M13 particles.[41] From this, we can say that the actual reaction yield is 8% (the yield calculation can be seen in appendix 7.3.1). Although still quite low, it was deemed as acceptable for purposes of LD detection, where it was envisaged that high levels of conjugation might not facilitate or improve LD changes in the alignment of M13-**R-For** conjugates after a PCR reaction, or even disrupt its behaviour prior to amplification. The conjugation levels were high enough and available in enough quantity that the conjugations could be used in excess as potential primers in a PCR.

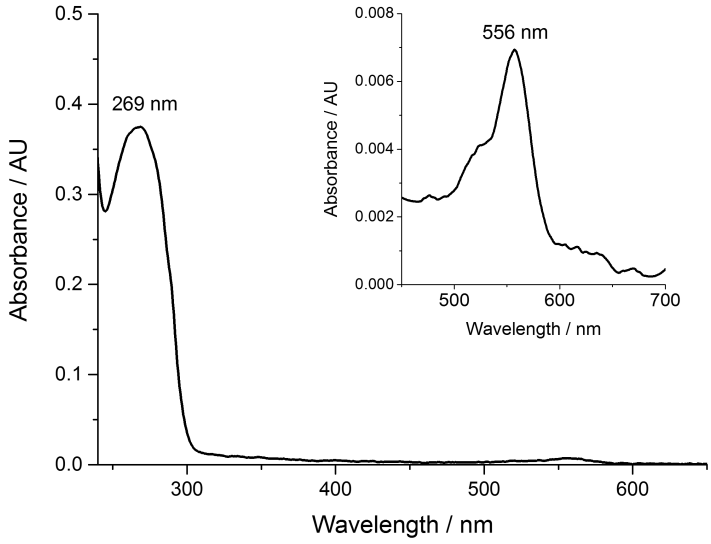


*Figure 6.12 - UV-vis spectrum of M13-**R-For**. Inset - zoomed in spectrum of the TAMRA absorbance at 556 nm.*

6.2.2.3 Effect on Linear Dichroism (N.S and JC-S)

In order to test the new conjugate for its use in LD based assays, it was compared with wtM13 (Figure 6.13) by alignment in shear flow and measurement by LD. Key to the success of any such assay is that the conjugate does not disrupt the alignment in flow drastically and therefore failing to produce a measurable signal. WtM13 (solid line) displayed a characteristic spectra identical to that shown previously.[39,41] Distinctive positive bands in the far UV-vis region (205 and 225 nm) and between 260 and 310 nm are the result of π-π* transitions of the tryptophan and tyrosine residues respectively. A positive band is observed due to the parallel orientation axis of these transitions with respect to the long molecular axis of the virus particle itself. The negative band observed between 240-260 nm is a result of the viral DNA bases encapsulated inside the particle. Encouragingly, these bands are only slightly perturbed by the conjugation of **R-For** to its coating (dashed line). This indicates that these short primers do not significantly alter the hydrodynamic behaviour of the particles in shear flow.

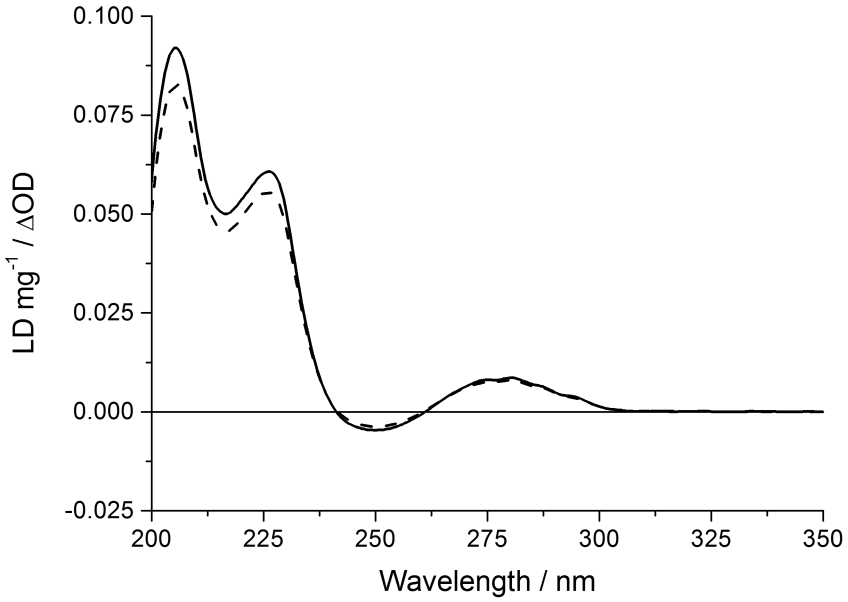


*Figure 6.13 - Comparison of the alignment of wtM13 (solid) and **M13-R-For** (dashed) in Couette flow monitored by LD. 50 mM potassium phosphate buffer pH 8.0,150 mM NaCl. Couette flow 3000 rpm. LD has been corrected for concentration and is shown as LD per mg of M13.*

6.2.3 Binding Studies

Although Lee *et al* had previously shown that short oligonucleotides appended to M13 could still participate in complementary base pairing using other methods,[19] it remained important to study the ability of these primers to bind DNA templates, to ensure the validity of this conjugate design. Modification of oligonucleotides (**R-For, F-Com, F-Ran**) with fluorophores allowed for fluorescence spectroscopy measurements of DNA binding, with a view to ensuring both complementary Watson-Crick hydrogen bonding was still possible, but also to ensure no non-specific interactions occur.

6.2.3.1 Anisotropy

Fluorescence anisotropy (FA) allows for the measurement of tumbling rates of molecules in solution which can be used to determine shape, molecular orientation and mobility as well as changes in molecular size.[50] FA works using standard fluorescence spectrometers, but differs in that it requires plane-polarised light in order for the experiments to work. The principle of FA is that if a molecule has an absorption transition moment in line with the polarisation plane of the excitation light source it will be selectively excited. Those that are not in the plane will remain in the ground state. As polarised light is absorbed and a fluorophore emits light, the light becomes less polarised. Anisotropy is a direct measurement in the difference between the polarisation of incident and emitted light. Where the anisotropy, $< r >$ can be defined as;[50]

Equation 6.2

$$< r > = \frac{I_{VV} - G I_{VH}}{I_{VV} + 2G I_{VH}}$$

The I_{XX} values represent the light intensities observed when the excitation and emission polarisers are mounted; both vertically (I_{VV}), both horizontally (I_{HH}), vertical (excitation) and horizontal (emission) (I_{VH}) and horizontal (excitation) and vertical (emission) (I_{HV}).G, the G-

factor, is a measure of the sensitivity of detection systems for horizontal and vertically polarised light, and is defined as;[50]

Equation 6.3

$$G = \frac{I_{HV}}{I_{HH}}$$

Modern spectrometers produce the anisotropy value automatically by applying the required polarisations, the value of which can be used to study binding effects of large biomolecules, in general as a molecule gets larger or binds a guest molecule its anisotropy increases (tumbles more slowly).

One of the requirements of accurately measuring fluorescence anisotropies is that the fluorophores cannot partake in resonance energy transfer (RET). For molecules that can do this, a reduction in the anisotropy values results due to depolarisation of the light. Therefore it was important to assess whether our oligonucleotides designed for studying the binding events were capable of RET, especially given that rhodamine (TAMRA) and fluorescein (6-FAM) derivatives are a well-known FRET pair, albeit when in a closer proximity to each other. Figure 6.14 below shows the emission spectra of **F-Com** when excited at 495 nm both for the unconjugated oligonucleotides and for the addition of M13-**R-For**. Notably neither system showed significant FRET peaks (a growing peak at 580 nm would be expected), suggesting that the 18 b.p separation, and perhaps location in the helical twist of the duplex is such that FRET cannot occur. Previously Clegg *et al.* have shown a system where the FRET distance is too large when separated by 16 or more base pairs for efficient energy transfer which coincides well with our findings.[51]

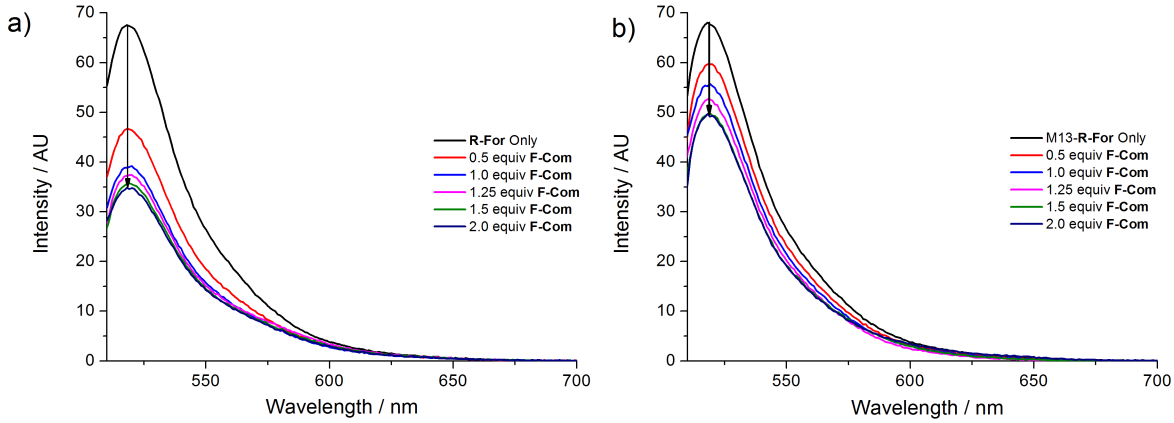


*Figure 6.14 - Fluorescence spectra of a) R-For and b) M13-**R-For** upon addition of complementary **F-Com** (0-2 equivalents) showing the absence of FRET behaviour. 0.25 μM DNA, 150 mM NaCl, 50 mM potassium phosphate buffer pH 8.0.*

Given that these systems were suitable for anisotropy experiments, the binding behaviour was investigated. Firstly, measurement of the individual counter-parts gave further confirmation of successful conjugation, where the FA of rhodamine (ex. 556 nm, em. 580 nm is higher (0.128 ± 0.0006) when appended to M13 than when free (0.061 ± 0.002). Figure 6.15 shows the FA changes when wtM13/M13-**R-For** conjugate is added to strands of fluorescein labelled oligonucleotides (**F-Com** and **F-Ran**). Addition of wtM13, which displays no specific binding sites for the DNA strands, to **F-Com** showed negligible changes in FA as expected, and confirmed there is no non-specific binding between the peptide and DNA constituents. Equally the addition of M13-**R-For** to a non-complementary **(F-Ran**) sequence also showed no change in the anisotropy, showing that non-specific DNA-DNA interactions are also absent.

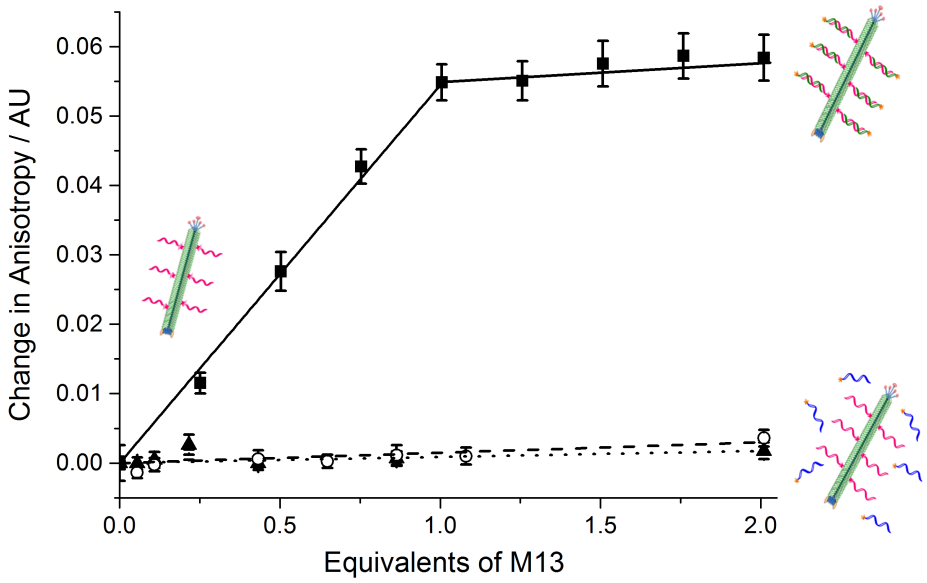


*Figure 6.15 - Fluorescence anisotropy changes of **F-Ran** and **F-Com** upon addition of M13-DNA conjugates and wtM13. Squares - F-Com + M13-**R-For**, circles - **F-Ran** + M13-**R-For**, triangles - **F-Com** + wtM13. 0.25 μM DNA, 150 mM NaCl, 50 mM potassium phosphate buffer, pH 8.0.*

Importantly, the addition of M13-**R-For** to its complement **F-Com** causes a dramatic increase in the FA, which plateaus at *ca*. 1 molar equivalent, as expected for DNA-DNA duplexes of this design. The anisotropy changes, although large, are not as great as those observed for M13-**R-For** alone. This is not unexpected. The fluorescence lifetime of the two fluorophores (6-FAM = 6 ns and TAMRA = 2 ns) are not long enough to be able to report on such large changes in molecular motion, in fact these liftetimes are only suitable for monitoring mass changes of 0-100 kDa.[50] The changes observed upon addition of **M13-R-For** are therefore attributed to contributions from both the change in size but also to the increased rigidity of the local environment around fluorescein. Hence, there follows a reduction in motion of the fluorophore and an increase in anisotropy. For comparison, the same trends were observed upon adding free oligonucleotide (**R-For**) to complementary and non-complementary targets as shown in Figure 6.16. As a whole, these results indicate that M13 conjugation does not affect the primers ability to form complementary sequence specific duplexes.

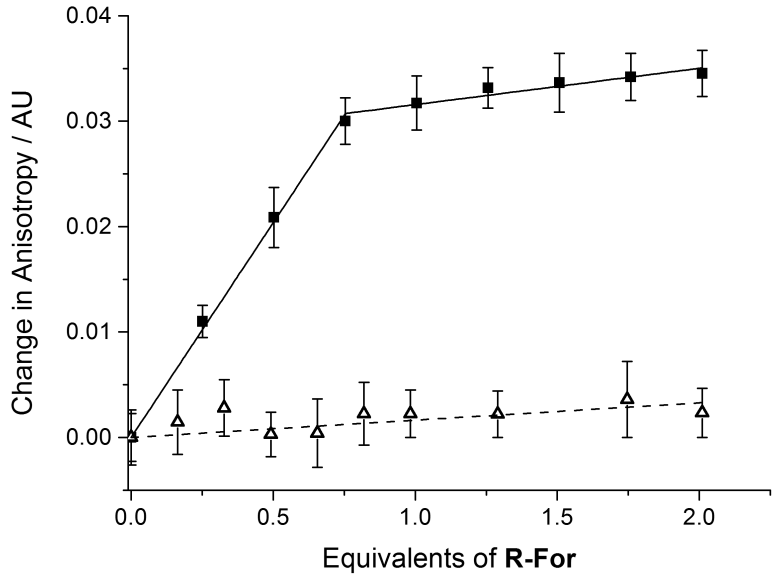


*Figure 6.16 - Change in fluorescence anisotropy observed for fluorescently labelled oligonucleotides; squares – **R-For** + F-Com and triangles – **R-For** + **F-Ran**. 0.25 μM DNA, 150 mM NaCl, 100 mM potassium phosphate buffer pH 8.0*

6.2.4 Polymerase Chain Reaction

Once the binding ability had been investigated, it was essential to our aims that the M13-DNA conjugate could partake in a PCR. Imperative to the workings of a PCR amplification is the denaturation step, which is usually carried out at temperatures in excess of 90 °C. Generally, proteins that are exposed to prolonged elevated temperatures, like in PCR, will begin to denature or aggregate and hence lose both function and structure. Therefore when these temperatures are then applied to M13, it may result in a misalignment and no LD signal. As such, the thermal stability of M13 was investigated, as described below.

6.2.4.1 Bacteriophage Stability (N.S and J.C-S)

LD was extremely useful in assessing directly the effect of temperature on M13 stability and hence alignment. Figure 6.17 below shows the effects of incubating M13 for 1 minute at a range of temperatures between 25 and 100 °C before measuring their LD at 3000 rpm. The data showed that it is not until temperatures are in excess of 86 °C that M13 alignment is significantly reduced.

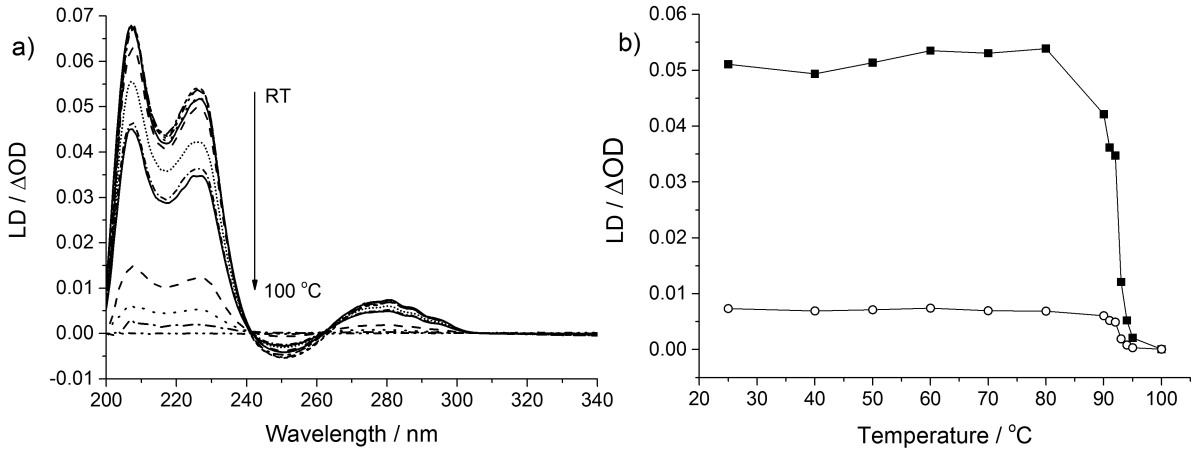


*Figure 6.17 - a) LD spectra of M13-**R-For** after being incubated under increasing temperature (25-100 °C) for 1 minute. b) Temperature dependence on electronic π-π* (225 nm) and aromatic transitions (280 nm) after incubation for 1 minute at increasing temperatures. M13 concentration = 0.11 mg/ml*

Additionally, subjection of M13 to 35 simulated PCR cycles at different temperatures indicated the same trend, in which the M13 alignment and hence the LD signal observed (Figure 6.18) drops off dramatically above 86 °C. Simulated PCR cycles consisted of a denaturation step, at the chosen temperature, annealing at 55 °C and extension at 72 °C (optimal for *Taq* polymerase) reflecting a typical *Taq* catalysed PCR amplification. Again this indicates that the stability of M13 is considerably diminished above 86 °C, where when the annealing temperature is increased to 92 °C, the LD signal diminishes by 99.5% compared to only 9% at 86 °C. From these sets of data it is clear that carrying out PCR above 86 °C would not be feasible for the assay and as such, PCR reactions should be tested at this lower temperature to test the validity of using *Taq* polymerase or the conjugates as PCR reagents.

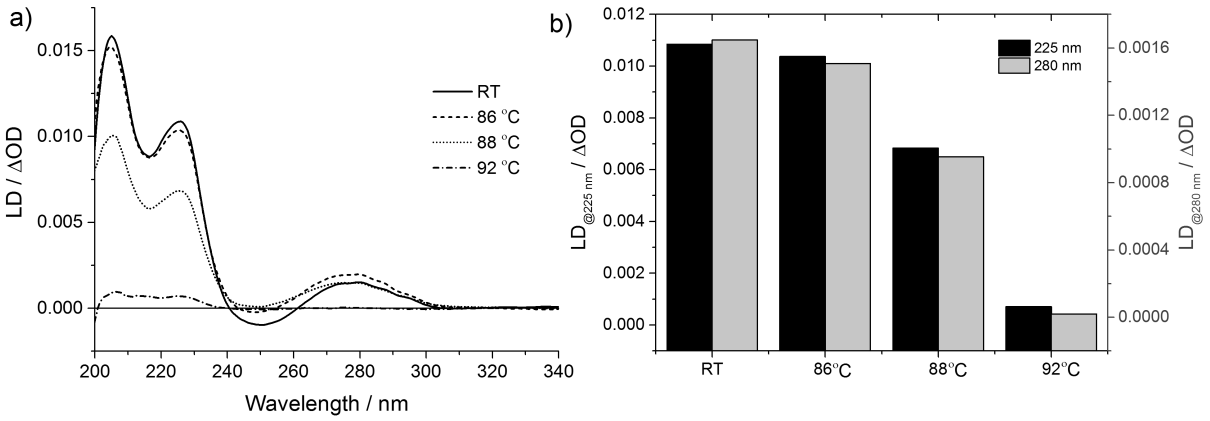


Figure 6.18 – a) LD spectra of M13 following 35 PCR cycles at elevated temperatures, b) bar chart showing the LD signal observed at 280 nm following 35 PCR cycles at varying temperatures. M13 concentration = 0.024 mg/ml.

6.2.4.2 PCR using M13-DNA conjugates (R.P-G and J.C-S)

Figure 6.19 shows a series of PCR reactions used to test the ability of a standard PCR amplification (Lane 2) to proceed at a lower temperature of 86 °C (usually 94 °C). It was anticipated that this would not be problematic, given that the predicted T_m values for the two primers are well below this temperature. Furthermore, previous work by Nadja Steinke had shown PCR amplification of *ampR* at 86 °C could be aided by the addition of ethylene glycol to the reaction mixture (see Appendix 7.3.2).[52] The ability of M13-**R-For** conjugates to amplify the ampR template (1 kbp in length) was also tested, by adding equal amounts of M13-**R-For** (based on 8 oligonucleotides per bacteriophage) to the reaction mixture and replacing the standard forward primer (Lane 3). The agarose gel obtained showed some interesting results, in which there was clear evidence for the PCR reaction using standard primers still being successful at slightly lower temperatures. Of most interest however was that the M13-**R-For** primer was also able to amplify the template as shown by the observation of a 1 kbp band. This band was also of significantly greater intensity, which may give some indication to the enhancement of PCR via this remarkable bio-nanoparticle. Evidence for these kinds of enhancements using nanoparticles has been observed previously and from this has grown a newly defined field of PCR technology, nanoPCR.[53]

Previous work has been restricted to nanoparticle suspensions involving AuNP,[54,55] quantum nanodots[56] and carbon nanotubes, [57] which have been able to enhance efficiency and specificity of PCR reactions. However it is believed this is the first example of PCR taking place on the surface of a naturally occurring bio-nanoscaffold.

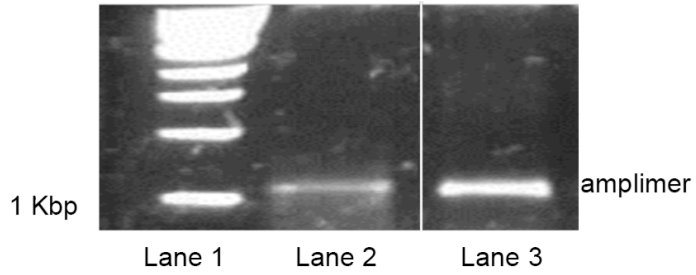


Figure 6.19 - Agarose gel showing the successful amplification of ampR at following 35 PCRcycles with a denaturation temperature of 86 °C. Lane 1 - DNA Ladder, lane 2 - amplification using standard forward and reverse primers, lane 3 - amplification following use of M13-R-For in place of standard forward primer R-For.

6.2.4.3 Linear Dichroism Studies (R.P-G and J.C-S)

The central aim of this study was to develop a M13-DNA conjugate that could participate in biosensing of pathogenic DNA via the amplification of genetic material of interest and subsequent detection of any such amplification by LD spectroscopy. Therefore for this to work, the amplimers produced, which are attached to the M13 particle, must disrupt the alignment enough to observe a change in the LD spectrum. From the number of PCR cycles and hence amplification time it was assumed that the PCR had reached its exponential plateau (the point at which all primers have been consumed) and therefore each M13-**R-For** has eight copies of the 1 kbp *ampR* amplimer attached. Each amplimer is approximately 0.33 μm in length (based on a 3.3 Å distance between base pairs), which compared with the M13 particle (0.88 μm), is still relatively short. The question remained as to whether the accumulation of hydrodynamic drag enforced by each amplimer was enough to cause a change in the LD.

The experiments included a number of important controls to ensure that any observed LD changes were the result of amplification and not due to false positives. A PCR carried out in the presence of *ampR* template with no PCR cycles, where it would be expected that amplification would not take place, is shown in Figure 6.20 (dashed black line) and displays the characteristic M13 transitions. For clarity, the LD spectrum (dashed green line) of a standard PCR reaction (Lane 2, Figure 6.19) following 35 PCR cycles was also probed, which displayed a negligible LD signal. This is important because it has previously been shown that DNA duplexes, albeit at much higher concentrations produce a measurable LD signal.[58] It was important to verify this to ensure that the LD spectrum observed for M13-**R-For** conjugates were a result of M13 alignment and not the amplimers themselves. The information provided by these controls allows for the unambiguous determination of the amplimer effects on M13 alignment.

A PCR reaction using M13-**R-For** in the absence of the *ampR* produces a characteristic LD spectrum of M13 (positive bands at 205, 225 and 280 nm, negative band at 254 nm). As expected, in the absence of appended amplimers, the M13-**R-For** conjugate remains able to align in its normal manner. However, following a successful amplification in the presence of the template, the appended amplimers are able to disrupt the alignment of the fibrous M13 particle beyond the maximum change that might be expected (i.e. LD becomes zero), instead producing a negative signal. The shape of this signal appears to be that of an inverted M13 spectrum, the usually positive bands now being negative and negative viral DNA transition approaching the positive. Therefore, instead of disrupting the alignment, it is possible that the amplimers are able to distort/change the orientation axis of the virus particle and hence lead to signal inversion.

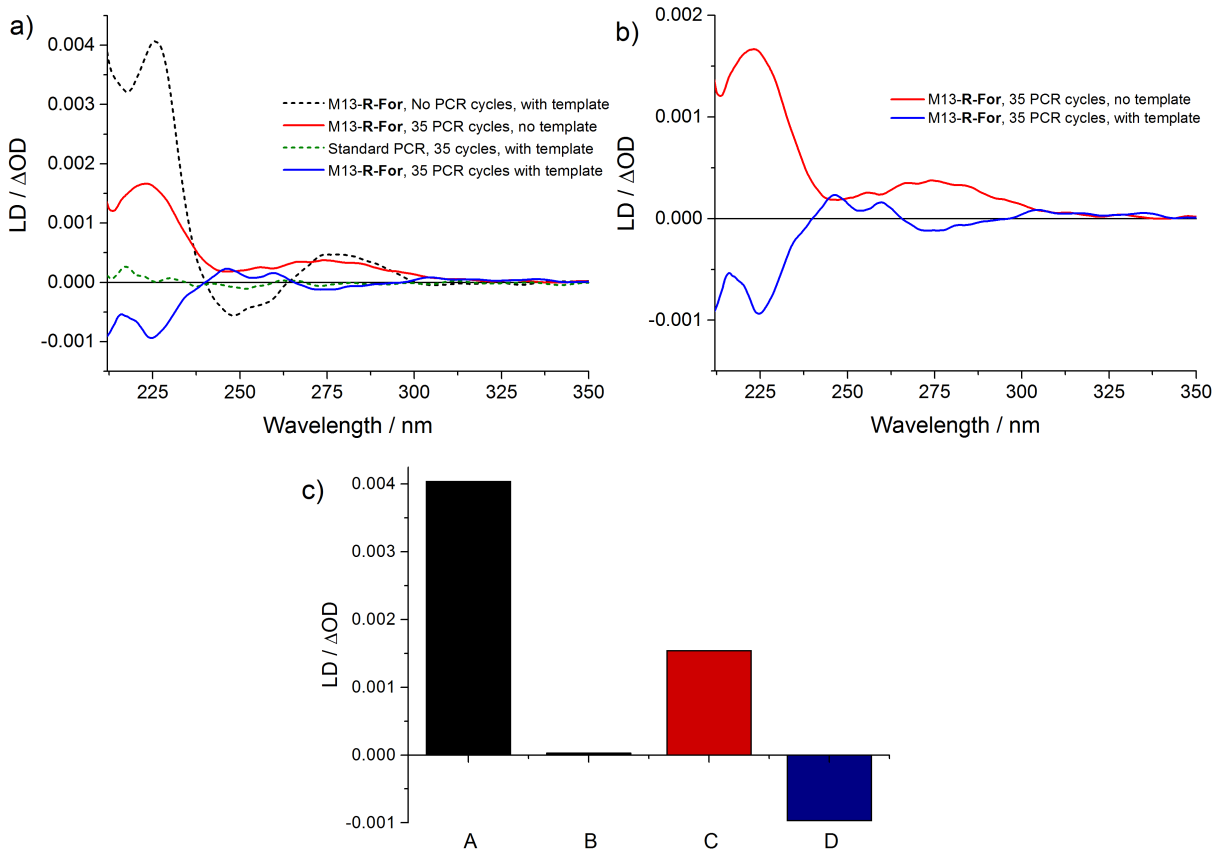


Figure 6.20 - a,b) LD spectra of the products of PCR reactions (86ºC denaturing temperature) in the absence and presence of target pathogenic DNA. c) bar chart depicting the LD signals at 225 nm for M13-R-For in the presence and absence of pathogenic DNA following 35 PCR cycles. A) M13-R-For, no cycles, B) R-For, 35 cycles, C) M13-R-For, 35 PCR cycles in the absence of ampR, D) M13-R-For, 35 PCR cycles in the presence of ampR.

The question as to why an inversion, rather than a signal reduction is observed, where the latter is usually observed as a result of disruption in alignment, can be addressed by considering the origins of LD signals. Clack *et al*[39] found that the LD signals are produced by the long axis alpha-helical nature of the major pVIII protein that constitutes the viral coating. The orientation of this protein with respect to the major orientation axis of the protein is *ca.* 20°. The sign and intensity of the signals produced by electronic transitions is dependent on the angle of that transition with respect to the alignment axis by a Cos^2 function. This means that, if the transition is aligned 54.7° from that axis, it will exhibit no LD signal, even if the

particle itself is aligned perfectly, a magic angle as such. However, if this angle is less than this so-called magic angle, the signal becomes positive, as observed in Figure 6.20. The pVIII protein starts with an angle of 20° in M13-**R-For** prior to amplification, which means a small misalignment of the particle of 34.7° would be required before the signal begins to invert, as demonstrated in Figure 6.21. Previously 'hairy' flagellates, which have long elongated bodies, have been shown to have dramatically different hydrodynamic properties to that of smooth exterior flagellates.[59,60] The hairs seemingly cause an inversion of the axial drag and hence alignment which alters the flow direction. The same concept can be used to explain the behaviour seen with the post-PCR products. M13 particles in shear flow align along two main axes (azimuthal and axial), the biased distribution towards particles pointing in the azimuthal direction causes the positive LD signals seen (Figure 9 a,b).[40] However, when the 0.33 µm amplimers are constructed on the surface of M13, particles aligned in the direction of flow (Figure 9 c) are disrupted as a result of increased drag. Those aligned axially (Figure 9 d) are stabilised which in turn causes a bias and hence deviation of the main electronic transitions enough to invert the signal.

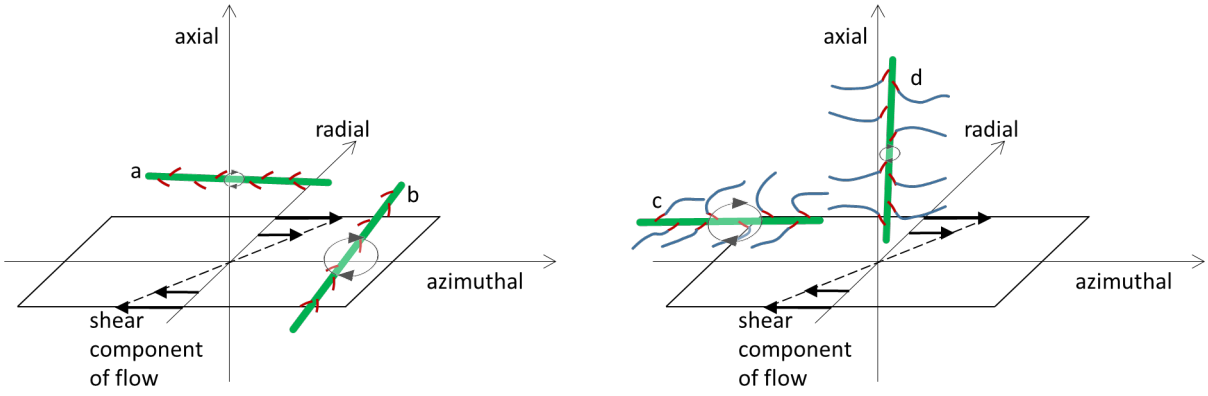


Figure 6.21 - Diagrammatic representation of M13 behaviour in Couette shear flow following PCR where a) and b) represent M13 in the absence and c) and d) in presence of ampR template. The diagram shows the principle axes of alignment. Figure constructed by Dr. David Smith.[1]

6.3 Conclusions

It has been shown for the first time that an M13 nanoscaffold can be assembled with DNA primers and can participate in an unprecedented example of nanoPCR on a virus surface. Additionally the presence of the amplimers, and hence confirmation of the presence of a pathogen template is detected by changes in the shear flow alignment of the M13-DNA scaffold, which results in an inversion of the LD signal. The detection of these amplimers is achieved within 1 minute of a PCR. It is hoped that the inversion property will allow for the extension of the dynamic range of such an assay, and perhaps opens possibilities for use in collaboration with quantitative (qPCR) and real-time (RT-PCR) techniques.

Beneficial to the ease of use of any such assay is the ability to indiscriminately deduce its success, and rule out potential false positives. Signal OFF sensors, as exemplified by the immuno-M13 assay previously reported,[41] are limited in this regard as signal reductions could be caused by potential degradation or contamination, resulting in misinterpretation of assay results. With our system producing an inverted signal as a result of the PCR, it has potential to be well suited for avoiding the need for complex control experiments, meaning that a non-trained person in the field could easily assess it. Not only does this show potential for enhanced sensitivity, but also it increases the dynamic range capabilities of such assays.

This study has broadened our understanding of the behaviour of M13 in shear flow and its ability to produce an LD signal. Importantly the development of M13 as a fibrous, alignable bio-scaffold in biosensing alongside LD shows great promise for the development of simultaneous multi-modal assays that could be used to detect a specific pathogen, i.e. detecting antigen, DNA and metabolites of a pathogen at the same time using the same instrumentation.

6.4 Future Work

The development of M13 scaffolds for biosensing using LD has opened up many possibilities for the creation of a whole host of new biosensors that could be appended to M13. The simplicity of the LD technique and the ability to miniaturise the spectroscopic instrumentation make it important to continue the focus on its use. Future work should include the following;

- In relation to this project and PCR;
    - Optimisation and enhancement of DNA conjugation to M13. Is it possible to directly control the amount of DNA appended to M13 for other applications?
    - A more in depth study of the benefits of M13 based nanoPCR over standard PCR reagents and other nanoPCR techniques. Does it improve efficiency, selectivity and specificity?
    - Further investigation into the robustness of M13-**R-For** for applications in LD based sensing of pathogenic DNA.
        - Effects of contamination
        - Effect of varying concentration. i.e. in hot climates
    - Study of the detection limits (in terms of template concentration) and the dynamic range of the assay itself.
    - Probing the use of templated PCR to create designed, highly ordered nanoassemblies based on M13 for the creation of new biomaterials.
    - Miniaturisation of this assay into hand-held PCR-LD devices that could be used in the field, clinic or at home.
- Ideas for the future;
    - Pre-organised complementary M13-DNA conjugates that aggregate but upon binding target DNA, dissociate and produce an LD signal, a signal ON sensor.

- M13-DNA conjugates based around aptamers for sensing of small molecules i.e. pathogen metabolites.

6.5 References

1. J. Carr-Smith, R. Pacheco-Gomez, H. A. Little, M. R. Hicks, S. Sandhu, N. Steinke, D. J. Smith, A. Rodger, S. A. Goodchild, R. A. Lukaszewski, J. H. R. Tucker and T. R. Dafforn, *ACS Synth. Biol.*, 2015, doi: 10.1021/acssynbio.1025b00034.
2. K. Brown, *Science*, 2004, **305**, 1228-1229.
3. J. J. Gooding, *Anal. Chim. Acta*, 2006, **559**, 137-151.
4. J. C. Guatelli, T. R. Gingeras and D. D. Richman, *Clin. Microbiol. Rev.*, 1989, **2**, 217-226.
5. M. J. Espy, J. R. Uhl, L. M. Sloan, S. P. Buckwalter, M. F. Jones, E. A. Vetter, J. D. C. Yao, N. L. Wengenack, J. E. Rosenblatt, F. R. Cockerill and T. F. Smith, *Clin. Microbiol. Rev.*, 2006, **19**, 165-256.
6. P. M. Holland, R. D. Abramson, R. Watson and D. H. Gelfand, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 7276-7280.
7. C. A. Heid, J. Stevens, K. J. Livak and P. M. Williams, *Genome Res.*, 1996, **6**, 986-994.
8. I. M. Mackay, K. E. Arden and A. Nitsche, *Nucleic Acids Res.*, 2002, **30**, 1292-1305.
9. Z. Dogic and S. Fraden, *Phys. Rev. Lett.*, 1997, **78**, 2417-2420.
10. D. Ghosh, A. F. Bagley, Y. J. Na, M. J. Birrer, S. N. Bhatia and A. M. Belcher, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13948-13953.
11. H. Yi, D. Ghosh, M.-H. Ham, J. Qi, P. W. Barone, M. S. Strano and A. M. Belcher, *Nano Lett.*, 2012, **12**, 1176-1183.
12. D. Ghosh, A. G. Kohli, F. Moser, D. Endy and A. M. Belcher, *ACS Synth. Biol.*, 2012, **1**, 576-582.
13. D. Oh, J. Qi, B. Han, G. Zhang, T. J. Carney, J. Ohmura, Y. Zhang, Y. Shao-Horn and A. M. Belcher, *Nano Lett.*, 2014, **14**, 4837-4845.
14. K. T. Nam, *Science*, 2006, **312**, 885-888.
15. Y. S. Nam, H. Park, A. P. Magyar, D. S. Yun, T. S. Pollom and A. M. Belcher, *Nanoscale*, 2012, **4**, 3405-3409.
16. C. Mao, D. J. Solis, B. D. Reiss, S. T. Kottmann and R. Y. Sweeney, *Science*, 2004, **303**, 213-217.
17. Y. S. Nam, T. Shin, H. Park, A. P. Magyar, K. Choi, G. Fantner, K. A. Nelson and A. M. Belcher, *J. Am. Chem. Soc.*, 2010, **132**, 1462-1463.
18. M. Brasino, J. H. Lee and J. N. Cha, *Anal. Biochem.*, 2014, **470**, 7-13.
19. J. H. Lee, D. W. Domaille and J. N. Cha, *ACS Nano*, 2012, **6**, 5621-5626.

20. B. Nordén, A. Rodger and T. Dafforn, *Linear Dichroism and Circular Dichroism*, Royal Society of Chemistry, 2010.
21. Y. Matsuoka and B. Nordén, *Chem. Phys. Lett.*, 1982, **85**, 302-306.
22. K. Razmkhah, H. Little, S. Sandhu, T. R. Dafforn and A. Rodger, *RSC Advances*, 2014, **4**, 37510-37515.
23. Y. Matsuoka and B. Nordén, *J. Phys. Chem.*, 1982, **86**, 1378-1386.
24. C. N. Rafferty and R. K. Clayton, *Biochim. Biophys. Acta*, 1978, **502**, 51-60.
25. J. Bolt and K. Sauer, *Biochim. Biophys. Acta*, 1979, **546**, 54-63.
26. M. R. Hicks, J. Kowałski and A. Rodger, *Chem. Soc. Rev.*, 2010, **39**, 3380-3393.
27. B. Nordén and F. Tjerneld, *Biophys. Chem.*, 1976, **4**, 191-198.
28. M. Jonsson, B. Åkerman and B. Nordén, *Biopolymers*, 1988, **27**, 381-414.
29. R. Marrington, E. Small, A. Rodger, T. R. Dafforn and S. G. Addinall, *J. Biol. Chem.*, 2004, **279**, 48821-48829.
30. R. Marrington, M. Seymour and A. Rodger, *Chirality*, 2006, **18**, 680-690.
31. K. E. Marshall, M. R. Hicks, T. L. Williams, S. V. Hoffmann, A. Rodger, T. R. Dafforn and L. C. Serpell, *Biophys. J.*, 2010, **98**, 330-338.
32. A. Rodger, R. Marrington, M. A. Geeves, M. Hicks, L. de Alwis, D. J. Halsall and T. R. Dafforn, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3161-3171.
33. K. Frykholm, K. Morimatsu and B. Nordén, *Biochemistry*, 2006, **45**, 11172-11178.
34. J. Rajendra and A. Rodger, *Chem. Eur. J*, 2005, **11**, 4841-4847.
35. X. Li, *Nucleic Acids Res.*, 2006, **34**, 3670-3676.
36. M. Ardhammar, N. Mikati and B. Nordén, *J. Am. Chem. Soc.*, 1998, **120**, 9957-9958.
37. M. R. Hicks, T. R. Dafforn, A. Damianoglou, P. Wormell, A. Rodger and S. V. Hoffmann, *Analyst*, 2009, **134**, 1623-1628.
38. A. Damianoglou, A. Rodger, C. Pridmore, T. R. Dafforn, J. A. Mosely, J. M. Sanderson and M. R. Hicks, *Protein Peptide Lett.*, 2010, **17**, 1351-1362.
39. B. A. Clack and D. M. Gray, *Biopolymers*, 1992, **32**, 795-810.
40. J. R. A. McLachlan, D. J. Smith, N. P. Chmel and A. Rodger, *Soft Matter*, 2013, **9**, 4977-4984.
41. R. Pacheco-Gómez, J. Kraemer, S. Stokoe, H. J. England, C. W. Penn, E. Stanley, A. Rodger, J. Ward, M. R. Hicks and T. R. Dafforn, *Anal. Chem.*, 2012, **84**, 91-97.
42. Kary B. Mullis - Nobel Lecture: The Polymerase Chain Reaction http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1993/mullis-lecture.html, Accessed 12 Jan, 2015.
43. K. Kleppe, E. Ohtsuka, R. Kleppe and I. Molineux, *J. Mol. Biol.*, 1971, **56**, 341-361.

44. R. Saiki, S. Scharf, F. Faloona, K. Mullis, G. Horn, H. Erlich and N. Arnheim, *Science*, 1985, **230**, 1350-1354.
45. R. Saiki, D. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis and H. Erlich, *Science*, 1988, **239**, 487-491.
46. J. G. Sutcliffe, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 3737-3741.
47. S. Sandhu, PhD Thesis, University of Birmingham, 2015.
48. K. Li, Y. Chen, S. Li, H. G. Nguyen, Z. Niu, S. You, C. M. Mello, X. Lu and Q. Wang, *Bioconj. Chem.*, 2010, **21**, 1369-1377.
49. D. W. Domaille, J. H. Lee and J. N. Cha, *Chem. Commun.*, 2013, **49**, 1759-1761.
50. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 3rd edn., 2007.
51. C. Gohlke, A. I. Murchie, D. M. Lilley and R. M. Clegg, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 11660-11664.
52. N. Steinke, Master's Dissertation, University of Birmingham, 2012.
53. D. Pan, Y. Wen, L. Mi, C. Fan and J. Hu, *Curr. Org. Chem.*, 2011, **15**, 486-497.
54. H. Li, J. Huang, J. Lv, H. An, X. Zhang, Z. Zhang, C. Fan and J. Hu, *Angew. Chem. Int. Ed.*, 2005, **44**, 5100-5103.
55. Y. Zhao, L. Xu, H. Kuang, L. Wang and C. Xu, *J. Mater. Chem.*, 2012, **22**, 5574-5580.
56. L. Ma, S. He, J. Huang, L. Cao, F. Yang and L. Li, *Biochimie*, 2009, **91**, 969-973.
57. X. Cao, X. Shi, W. Yang, X. Zhang, C. Fan and J. Hu, *Analyst*, 2008, **134**, 87-92.
58. D. J. Halsall, A. Rodger and T. R. Dafforn, *Chem. Commun.*, 2001, 2410-2411.
59. M. E. Holwill and M. A. Sleigh, *J. Exp. Biol.*, 1967, **47**, 267-276.
60. T. L. Jahn, M. D. Landman and J. R. Fonseca, *J. Protozool.*, 1964, **11**, 291-296.

Chapter 7

Experimental

7.1 Synthesis

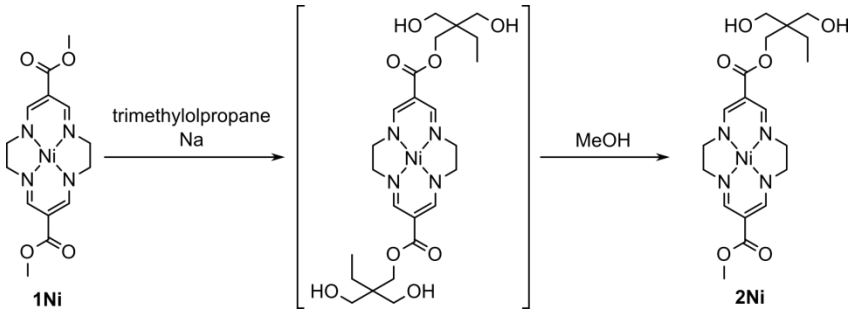
7.1.1 Synthesis of Ferrocene Phosphoramidites

The synthesis of non-commercially available phosphoramidites discussed within this thesis was carried out by the Dr. Huy. Van Nguyen, Jonathan Kedge (FcNA –Chapters 3 and 4) The full procedures for the synthesis of these compounds are detailed in the published results described in the relevant chapters.

7.1.2 Synthesis of Macrocyclic Phosphoramidites

The synthesis of all macrocyclic phosphoramidites was carried out by Dr. Jarsoław Kowalski (Institute of Organic Chemistry, Polish Academy of Science). The full synthetic experimental details are included below as carried out and written by Dr. Kowalksi.

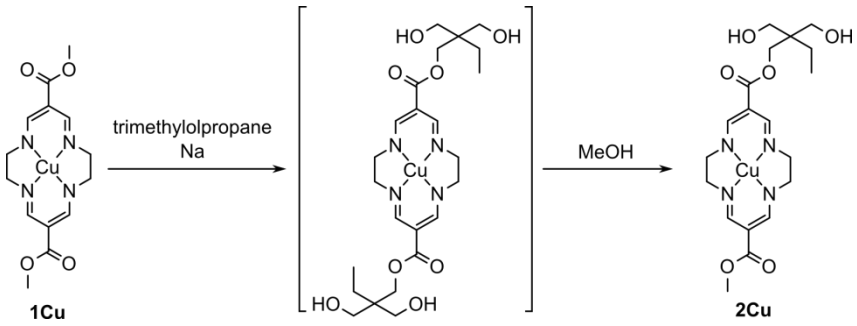
Synthesis of 6-carbo[2,2-bis(hydroxymethyl)]butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$nickel(II) (**2Ni**)



21.2 g (0.16 mol) of trimethylolpropane (1,1,1-tris(hydroxymethyl)propane) was heated (90 °C) in a round-bottomed flask until all of the alcohol had melted. Then 0.730 g (2 mmol) of **1Ni** and ca 0.5 g of sodium were added while stirring. The mixture was stirred at 90 °C for 24 h and monitored by TLC (silica gel, 4 % of MeOH in DCM). After the full transesterification was complete, the heating bath temperature was reduced to 80 °C and another 0.5 g of sodium, followed by 5 mL of methanol, were added. The reaction was followed carefully by TLC. When the spot of monotransesterified product (**2Ni**) on the TLC had reached its maximum intensity the reaction was cooled down to r.t. and, after making sure all of the sodium was dissolved, 75 mL of water was added. The mixture was then poured into a flask containing 250 mL of ice-water and left for several hours. The precipitate was filtered off and

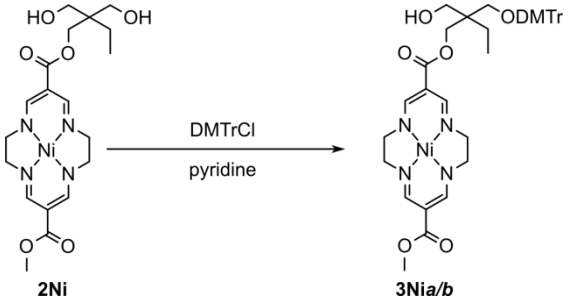
washed with water. After drying off in vacuo the solid was dissolved in DCM and loaded onto a silica gel 60 column. DCM containing 3% methanol was used as an eluent and the product was collected as the second orange band. The fractions were evaporated to dryness and the product was recrystallised from DCM in a freezer. Filtering off and drying in vacuo afforded **2Ni** as orange crystalline solid with a yield of 0.350 g (0.741 mmol, 37 %). ^{1}H NMR ($CDCl_3$, 200 MHz, ppm): 0.89 (d, J = 7.5 Hz, 3H, $–CH_3$), 1.22 (q, 2H, J = 7.5 $–CH_2–$ α to $–CH_3$), 2.53 (br s, 2H, –OH assoc. with H_2O), 3.39 (br m, 8H, NCH_2CH_2N), 3.54 (m, 4H, $–CH_2–$ α to –OH), 3.73 (s, 3H, $–OCH_3$), 4.34 (s, 2H, $–CH_2OCO–$), 7.81 (s, 4H, –CH=N); ^{13}C NMR ($CDCl_3$, 50 MHz, ppm): 7.9 ($–CH_3$), 23.3 ($–CH_2–$ α to $–CH_3$), 43.5 (4° C β to –OH), 51.2 ($–OCH_3$), 59.1 and 59.3 (NCH_2CH_2N), 63.4 ($–CH_2–$ α to –OCO–), 66.5 ($–CH_2–$OH), 97.4 (ring *meso* –C=), 155.4 (br, ring –C=N), 167.5 and 168.0 (C=O); MS (FD, CH_2Cl_2, m/z): 466.1 ($[C_{19}H_{28}N_4O_6Ni]^+$); IR (nujol, cm^{-1}): 3379 m, 1671 s, 1646 s, 1595 s, 1539 s, 1275 s, 1121 s, 762 s; Elemental analysis calcd for $C_{19}H_{28}N_4O_6Ni$ (467.1): C 48.85, H 6.04, N 11.99; found: C 48.90, H 6.02, N 11.82.

Synthesis of 6-carbo[2,2-bis(hydroxymethyl)]butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$copper(II) (**2Cu**)



The product was synthesised according to the method described above, using 0.740 g (2 mmol) of the copper(II) ester **1Cu**. The product was isolated as red crystals with a yield of 0.402 g (0.85 mmol, 35 %). Recrystallisation from dichloromethane upon slow evaporation of the solvent afforded red crystals suitable for single crystal X-ray structural measurements. MS (FD, CH_2Cl_2, m/z): 471.1 ($[C_{19}H_{28}N_4O_6Cu]^+$); IR (nujol, cm^{-1}): 3379 m; 2725 m; 1674 s; 1649 s; 1594 s; 1547 s; 1265 s; 1121 s; 764 s; Elemental analysis calcd for $C_{19}H_{28}N_4O_6Cu$ (472.0): C 48.35, H 5.98, N 11.87; found: C 48.38, H 5.88, N 11.74.

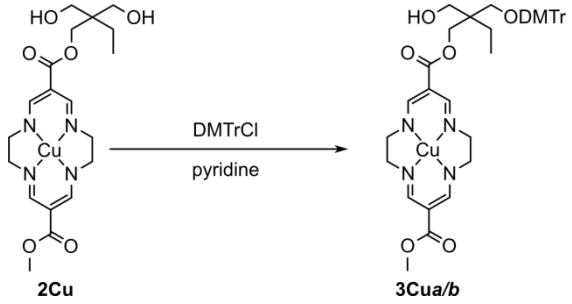
Synthesis of 6-carbo[2-(hydroxymethyl)-2-(4,4'-dimethoxytriphenylmethoxy)methyl]butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$nickel(II) (**3Ni*a/b***)



2Ni **3Ni*a/b***

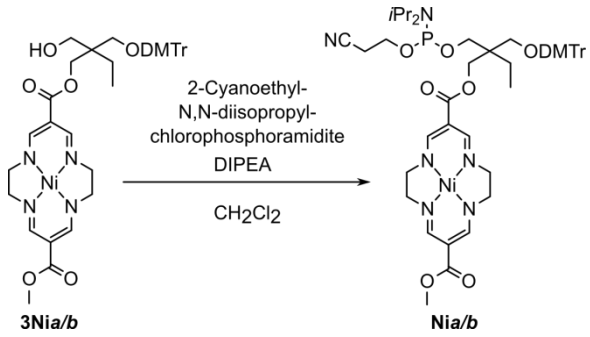
0.310 g (0.66 mmol) of **2Ni** was dissolved in 2.5 mL of dry pyridine. Subsequently, a solution of 0.119 g (0.35 mmol) of 4,4'-dimethoxytrityl chloride (DMTrCl) in 1.5 mL of dry pyridine was added dropwise (syringe pump, 2 mL/h) with stirring. After 0.5 h the addition was complete, the solvent was evaporated and 2 mL of toluene was added. Again, the solvent was evaporated and the process was repeated several times in order to remove pyridine. The solid was dissolved in a small amount of DCM, loaded onto a silica gel column, and eluted with 1:1 ethyl acetate / hexane solution with 0.25 % of Et_3N. The major coloured (orange) fraction was collected and evaporated to dryness. The remaining amine was then removed under vacuum, which gave 0.360 g (0.47 mmol, 70 %) of **3Ni*a/b*** as an orange foam. ^{1}H NMR (CD_3CN, 400 MHz, ppm): 0.68 (d, 3H, $-CCH_3$), 1.41 (m, 2H, $-CH_2-$Me), 2.95 (m, 2H, $-CH_2-$O–DMTr), 3.34 (s, 8H, NCH_2CH_2N), 3.39 (m, 2H, $-CH_2-$ α to –OH), 3.64 (s, 3H, $-OCOCH_3$), 3.78 (s, 6H, CH_3-O–Ph), 4.07 (m, 2H, $-CH_2-$OCO–), 6.82 (m, 4H, phenylene, *o* rel. to –OMe), 7.21 (m, 1H, phenyl, *p*), 7.26 (m, 2H, phenyl, *m*), 7.27 (m, 4H, phenylene, *m* rel. to –OMe), 7.41 (m, 2H, phenyl, *o*), 7.63 (s, 2H, –CH=N), 7.78 (s, 2H, –CH=N); MS (FD, CH_2Cl_2, m/z): 768.2 ($[C_{40}H_{46}N_4O_8Ni]^+$); Elemental analysis calcd for $C_{40}H_{46}N_4O_8Ni$ (769.5): C 62.43, H 6.03, N 7.28; found: C 62.52, H 5.97, N 7.19.

Synthesis of 6-carbo[2-(hydroxymethyl)-2-(4,4'-dimethoxytriphenylmethoxy)methyl]butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$copper(II) (**3Cu*a/b***)



The product was synthesised following the above procedure. 0.312 g (0.66 mmol) of **2Cu** was used as a substrate, and the product (pale red foam) was obtained with a yield of 78 % (0.399 g, 0.52 mmol). MS (FD, CH_2Cl_2, m/z): 773.3 ($[C_{40}H_{46}N_4O_8Cu]^+$); Elemental analysis calcd for $C_{40}H_{46}N_4O_8Cu$ (774.4): C 62.04, H 5.99, N 7.24; found: C 62.09, H 5.92, N 7.14.

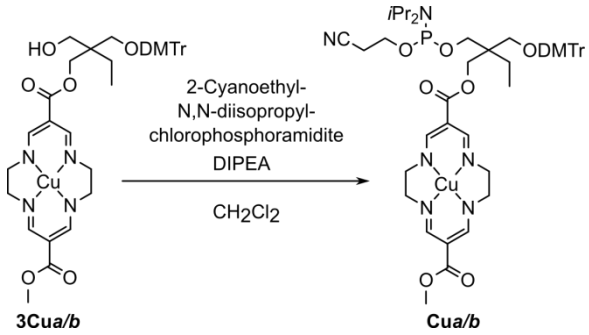
Synthesis of 6-carbo[2-(*O*-(2-cyanoethyl-*N*,*N*-diisopropylphosphoramidyl)methyl)-2-(4,4'-dimethoxytriphenylmethoxy)methyl]-butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)- $\kappa^4N^{1,4,8,11}$nickel(II) (**Ni*a/b***)



0.200 g (0.26 mmol) of **3Ni*a/b*** was dissolved in 5 mL of dry dichloromethane. Whilst stirring under argon atmosphere, 0.21 mL of *N*,*N*-diisopropylethylamine (DIPEA) and, subsequently, 0.07 mL (0.312 mmol) of 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, were added. The stirring was continued at room temperature for 3 h and 10 mL of ethyl acetate was added. The solution was washed twice with saturated aqueous $NaHCO_3$ solution and once with saturated NaCl(aq.). The organic layer was dried with Na_2SO_4 and filtered. The solvents were evaporated and the crude product was dissolved in 1:1 mixture of ethyl acetate and hexane (20 mL). The resulting solution was applied onto a silica gel 60 column and eluted with 2:1 ethyl acetate / hexane with an addition of 0.25 mL of Et_3N per

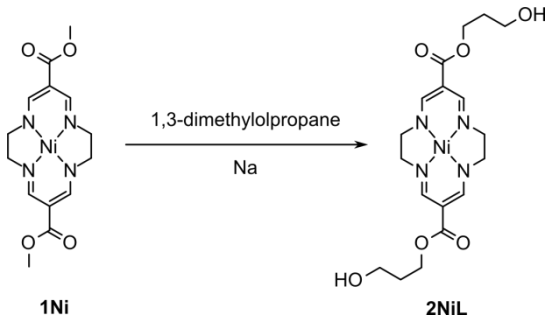
each 100 mL of the eluent. The orange fraction was collected and subsequently evaporated to give a pale orange foam. The solid was additionally dried in vacuo to obtain 0.230 g (0.24 mmol, 91 %) of **Ni*a/b***. ^{1}H NMR (CD_2Cl_2, 400 MHz, ppm): 0.74 (d, 3H, –CCH_3), 1.16 (m, 12H, –CH(CH_3)$_2$), 1.49 (m, 2H, –CH_2–Me), 2.66 (t, J = 6.4 Hz, 2H, –CH_2CN), 3.04 (m, 2H, –CH_2–O–DMTr), 3.36 (s, 8H, NCH_2CH_2N), 3.55 (m, 2H, –CHN(Me)$_2$), 3.68 (m, 2H, –OCH_2CH_2CN), 3.69 (s, 3H, CH_3OCO–), 3.81 (s, 6H, CH_3OPh), 3.85 (m, 2H, –CCH_2OP), 4.12 (m, 2H, –CH_2OCO–), 6.80 (m, 4H, phenylene, *o*), 7.22 (m, 1H, phenyl, *p*), 7.26 (m, 2H, phenyl, *m*), 7,31 (m, 4H, phenylene, *m*), 7,43 (m, 2H, phenyl, *o*), 7,66 (s, 2H, –CH=N), 7,82 (s, 2H, –CH≡N); ^{31}P NMR (CD_3CN, 121.5 MHz, ppm): 146.8 and 147.0 (two diastereomeric pairs of enantiomers); MS (FD, CH_2Cl_2, m/z): 968.4 ([$C_{49}H_{63}N_6O_9PNi$]$^+$).

Synthesis of 6-carbo[2-(*O*-(2-cyanoethyl-*N*,*N*-diisopropylphosphoramidyl)methyl)-2-(4,4'-dimethoxytriphenylmethoxy)methyl]-butoxy-13-carbomethoxy-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)- $\kappa^4N^{1,4,8,11}$copper(II) (**Cu*a/b***)



The synthesis of **Cu*a/b*** was carried out according to the procedure described for analogous nickel(II) complexes (**Ni*a/b***) shown above. 0.240 g (0.310 mmol) of **3Cu*a/b***, 0.250 mL of DIPEA and 0.083 mL (0.372 mmol) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite were used. The product (**Cu*a/b***) was obtained as pale red foam. The yield was 0.232 g (0.238 mmol, 77 %). MS (FD, CH_2Cl_2, m/z): 973.4 ([$C_{49}H_{63}N_6O_9PCu$]$^+$).

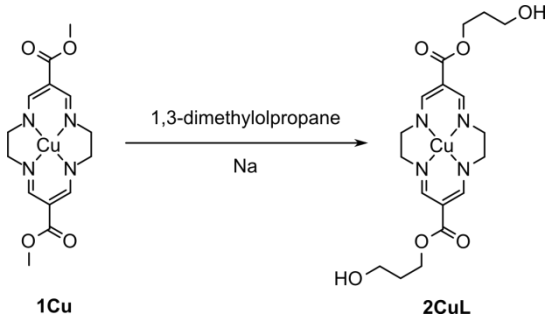
Synthesis of 6,13-bis[carbo(3-hydroxypropoxy)]-1,4,8,11-tetraazacyclotetradeca-4,6,11,3-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$nickel(II) (**2NiL**)



1Ni **2NiL**

The synthesis of **2NiL** was carried out according to our previously reported procedure[1,2] with the following modifications. The amount of sodium was increased 5 times (ca 10 equiv. with respect to the substrate). Moreover, the reaction was carried out for 48 h with the temperature lowered to 80 °C. The modifications allowed for increasing the yield to 92 %. The structure of the product was confirmed using NMR, ESI MS and elemental analysis. All the analytical results were found to be identical to those already published.

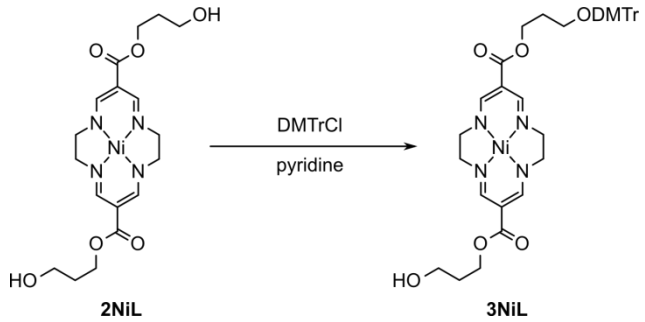
Synthesis of 6,13-bis[carbo(3-hydroxypropoxy)]-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)- $\kappa^4N^{1,4,8,11}$copper(II) (**2CuL**)



1Cu **2CuL**

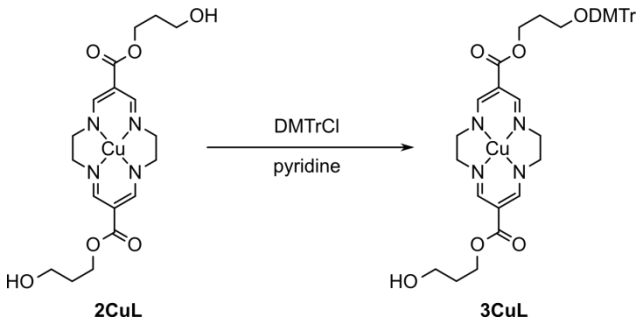
2CuL was synthesised following already published procedures[1,2] with modifications as described above (for **2NiL**). The yield increased to 96 % and all the analyses (ESI MS and elemental analysis) were identical to those found in literature. Recrystallisation from dichloromethane at 5 °C gave red crystals of a quality sufficient for singe crystal X-ray data collection.

Synthesis of 6-carbo[3-(4,4'-dimethoxytriphenylmethoxy)propoxy]-13-carbo(3-hydroxypropoxy)-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)- $\kappa^4N^{1,4,8,11}$nickel(II) (**3NiL**)



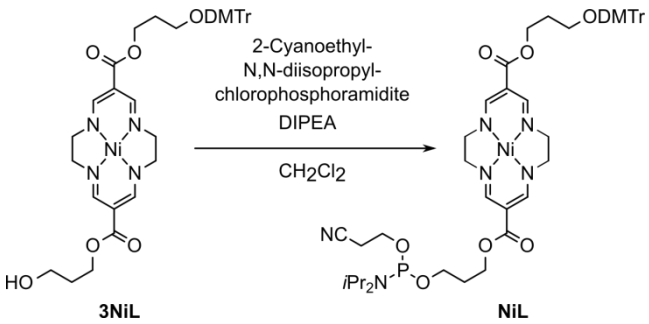
3NiL was obtained following the same procedure as described for **3Ni*a/b***, with larger excess of DMTrCl (1.5 eq.). The first fraction from a silica gel column was the ditritylated by-product (yield 28 %). **3NiL** was collected as the second fraction, and isolated as a pale orange foam with 52 % yield. ^{1}H NMR (CD_2Cl_2, 400 MHz, ppm): 1.89 (quint, J = 6.0 Hz, 2H. $-CH_2-$ β to –OH), 1.95 (quint, J = 6.1 Hz, 2H, $-CH_2-$ β to –ODMTr), 3.17 (t, J = 6.0 Hz, 2H, $-CH_2-$ α to –ODMTr), 3.35 (m, 8H, NCH_2CH_2N), 3.70 (t, J = 6.0 Hz, 2H, $-CH_2-$ α to –OH), 4.27 (t, J = 6.2 Hz, 2H, $-CH_2$ γ to –ODMtr), 4.33 (t, J = 6.0 Hz, 2H, $-CH_2$ γ to –OH), 6.82 (m, 4H, phenylene, *o* rel. to –OMe), 7.22 (m, 1H, phenyl, *p*), 7.27 (m, 2H, phenyl, *m*), 7.30 (m, 4H, phenylene, *m* rel. to –OMe), 7.41 (m, 2H, phenyl, *o*), 7.71 (s, 2H, –CH=N), 7.82 (s, 2H, –CH=N); MS (FD, CH_2Cl_2, m/z): 754.2 ($[C_{39}H_{44}N_4O_8Ni]^+$); Elemental analysis calcd for $C_{39}H_{44}N_4O_8Ni$ (755.5): C 62.00, H 5.87, N 7.77; found: C 61.89, H 5.95, N 7.70.

Synthesis of 6-carbo[3-(4,4'-dimethoxytriphenylmethoxy)propoxy]-13-carbo(3-hydroxypropoxy)-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)- $\kappa^4N^{1,4,8,11}$copper(II) (**3CuL**)



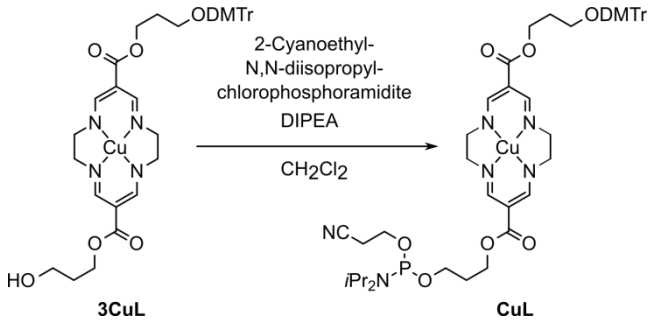
The product was prepared analogously to that described above for **3NiL**.The yield of two DMTr-functionalised side product was 53 %. **3CuL** was isolated as pale red foam with the yield of 46 %. MS (FD, CH_2Cl_2, m/z): 759.2 ($[C_{39}H_{44}N_4O_8Cu]^+$); Elemental analysis calcd for $C_{39}H_{44}N_4O_8Cu$ (760.3): C 61.61, H 5.83, N 7.37; found: 61.63, H 5.78, N 7.31.

Synthesis of 6-carbo[3-(4,4'-dimethoxytriphenylmethoxy)propoxy]-13-carbo[(2-cyanoethyl-N,N-diisopropylphosphoramidyloxy)propoxy]-1,4,8,11-tetraazacyclotetradeca-4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$nickel(II) (**NiL**)



The phosphitylation of **3NiL** was carried out in exactly the same way as for the obtaining of **Ni*a/b***. The product (**NiL**, as orange foam) yield was 85 %. ^{1}H NMR (CD_2Cl_2, 400 MHz, ppm): 1.19 (m, 12H, –CH(C*H*$_3$)$_2$), 1.94 (quint, 2H, J = 6.2 Hz, –CH_2– β to –ODMTr), 1.97 (quint, 2H, J = 6.3 Hz, –CH_2– β to –OP), 2.64 (t, J = 6.4 Hz, 2H, –CH_2CN), 3.17 (m, J = 6.0 Hz, 2H, –CH_2–ODMTr), 3.35 (m, 8H, NCH_2CH_2N), 3.62 (m, 2H, –CHN(Me)$_2$), 3.76 (m, 2H, –CH_2– α to –OP), 3.85 (m, 2H, –OC*H*$_2$CH$_2$CN), 4.22 (m, 2H, –CH_2–OCO– γ to –OP), 4.33 (t, J = 6.3 Hz, 2H, –CH_2–OCO– γ to –ODMTr), 6.83 (m, 4H, phenylene, *o* rel. to –OMe), 7.22 (m, 1H, phenyl, *p*), 7.28 (m, 2H, phenyl, *m*), 7.32 (m, 4H, phenylene, *m* rel. to –OMe), 7.44 (m, 2H, phenyl*, o*), 7.70 (s, 2H, –CH=N), 7.82 (s, 2H, –CH=N); ^{31}P NMR (CD_3CN, 121.5 MHz, ppm): 147.0; MS (FD, CH_2Cl_2, m/z): 954.4 ($[C_{48}H_{61}N_6O_9PNi]^+$).

Synthesis of 6-carbo[3-(4,4'-dimethoxytriphenylmethoxy)propoxy]-13-carbo[(2-cyanoethyl-N,N-diisopropylphosphoramidyloxy)propoxy]-1,4,8,11-tetraazacyclotetradeca 4,6,11,13-tetraenato(2-)-$\kappa^4N^{1,4,8,11}$copper(II) (**CuL**)



The product (**CuL**) was obtained following the same procedure as **Cu*a/b***, with the yield (red foam) of 80 %. MS (FD, CH_2Cl_2, m/z): 959.3 ([$C_{48}H_{61}N_6O_9PCu$]

7.2 Ferrocene Nucleic Acids and Macrocycles

7.2.1 Materials and Methods

All materials and reagents were purchased from either Sigma Aldrich (Gillingham, UK) or Fisher Scientific (Loughborough, UK) unless otherwise stated and used as provided without further purification. All DNA reagents were purchased from Link Technologies (Bellshill, UK) unless otherwise stated and used without further purification.

7.2.2 Automated Oligonucleotide Synthesis

All oligonucleotides were synthesised on an Applied Biosystems ABI 394 DNA/RNA synthesiser using standard phosphoramidite chemistry unless otherwise stated. Dry DNA synthesis grade dichloromethane and trichloroacetic acid in dichloromethane were obtained from Life Technologies (Paisley, UK). DNA grade dry acetonitrile was purchased from Rathburn Chemicals Ltd (Scotland, UK). Syntheses were generally carried out on a 1.0 μM scale. 0.5 M ETT in acetonitrile was used throughout as the activator and 0.02 M iodine in water solution was used for oxidation. Capping reagents were as follows; 10% acetic

anhydride in THF and 10% methylimidazole in THF for standard DNA synthesis. For ULTRA-mild synthesis a THF/pyridine/pac-anhydride (85:10:5) mixture was used in place of acetic anhydride. All phosphoramadites were dissolved to a concentration of 0.1 M in DNA synthesis grade anhydrous acetonitrile immediately before use. Coupling of standard phosphoramadites was for 25 s. Oligonucleotides were cleaved from the CPG solid supports by treatment with a 30% ammonia hydroxide solution for 90 minutes. Final deprotection of nucleobases was carried out by heating the oligonucleotides for 6 hours at 55 ^{o}C in the same ammonia hydroxide solution.

Modifications were made by the phosphoramadite method. This required the monomer to be pre-functionalized with a phosphoramadite group via an alcoholic functionality. For use on the DNA synthesiser, all synthesized phosphoramadites were dissolved in dry acetonitrile (100 mg ml^{-1}) and filtered through a 0.2 µm PTFE syringe filter (13 mm diameter, Whatman). The solvent was then removed under reduced pressure. The residues were then redissolved in anhydrous DCM and dispensed into pre-weighed vials fitted with rubber septa. These were dried over KOH under high vacuum overnight. The vials were then filled with argon and stored in a -20 ^{o}C freezer until use.

Unless otherwise stated, the coupling time was extended to that recommended by the manufacturer, or 10 mins in the case of phosphoramadites made by Tucker group, unless otherwise stated. Cleavage and deprotection conditions were followed according to manufacturer's instructions for purchased modifiers. For synthesised phosphoramadites, standard deprotection times were used throughout unless otherwise described.

7.2.3 Purification

Synthesised oligonucleotides were purified using Reversed-Phased High Performance Liquid Chromatography (RP-HPLC). A Dionex system with Summit P580 pump and Summit UVD 170s UV/VIS Multi-Channel Detector with prep flow cell fitted with a C18 column

(Phenomenex Clarity Oligo-RP columns, 150 mm x 10.0 mm 5 micron) were used throughout. Oligonucleotides were eluted with 0.1 M TEAA, pH 7.0 and acetonitrile gradients. Four different solvent systems were used throughout and were as follows; **Buffer A** - 5% acetonitrile in 0.1 M TEAA buffer pH 7.0; **Buffer B** – 15% acetonitrile in 0.1 M TEAA buffer pH 7.0; **C** - Acetonitrile; **Buffer D** – 0.1 M TEAA buffer pH 7.0. The elution was monitored using wavelengths of 260 and 210 nm. Major fractions were collected, dried down using a rotary evaporator and desalted using NAP-10 G25 Sephadex desalting columns (GE Healthcare). Purified oligonucleotides were subsequently characterized and their purity checked by analytical RP-HPLC using a Shimadzu Analytical HPLC system with a Phenomenex Clarity Oligo-RP analytical column (150 mm x 4.6 mm, 5 micron. The gradient system used for each oligonucleotide are detailed in appendix 1 and the systems are as follows:

- ThiolFerro: A two solvent system using Buffer D and MeCN gradients was used in the following time frame;
    - 5-20% MeCN: 35 minutes, 100% MeCN: 7 minutes, 5% MeCN: 8 minutes. Solvent rate: 3 ml per minute.
- Oligo 60: A three solvent system using Buffer A, Buffer D and MeCN gradients was used in the following time frames;
    - 30-50% Buffer B: 25 minutes, 50-100% Buffer B: 7 minutes, 100-0% Buffer B: 13 minutes, 0% Buffer B: 10 minutes, 30% Buffer B, 5 minutes
    - 0% MeCN: 35 minutes, 0-100% MeCN: 10 minutes, 100% MeCN: 10 minutes, 0% MeCN: 5 minutes

7.2.4 Mass Spectrometry

7.2.4.1 Oligonucleotide Characterisation

Mass spectra of purified oligonucleotides was carried out on either a Waters LCT MicroMass Electrospray Ionisation (ESI) Mass Spectrometer fitted with a time-of–flight (TOF) mass analyser or a Waters Synapt G2-S1 mass spectrometer fitted with an ESI ionisation source and a TOF mass analyser. Samples of 30 µL of ca. 100 µM concentration were first diluted to 500 µL using 50/50 mixture of methanol and water containing 1% TEA. Raw data was analysed and deconvoluted using Waters MaxEnt software. Predicted masses were obtained using OligoCalc an online web tool for predicting spectroscopic and mass properties of oligonucleotides.[3]

Details of the predicted and observed masses for all the oligonucleotides discussed within this thesis are included in the appendices.

7.2.4.2 Mercury Binding

50 µM of unmodified (sequence) or modified (sequence) and 50 µM of target strands (sequence descriptor) was dissolved in 50 µL of 50 mM NH_4OAc. To which a 1.2 equivalent excess of $Hg(ClO_4)_2$ was added. The solution was mixed thoroughly, heated to 85 $^{\circ}C$ for 5 minutes and slowly annealed overnight to room temperature. The samples were then diluted in MeOH/water (50:50) before direct infusion into the Synapt G2-S1 mass spectrometer and ionisation by electrospray with a capillary voltage of 2.58 keV a sampling cone voltage of 30 keV, source temperature of 100 ^{o}C and desolvation temperature of 250 $^{\circ}C$.

7.3 Optical Spectroscopy

7.3.1 UV-vis Spectroscopy

DNA, RNA and PNA concentrations were determined using a Shimadzu 1800 UV-Vis spectrophotomer. DNA molar extinction coefficients were calculated using the molar absorption coefficients for DNA predicted by the nearest neighbour calculation.[4,5] For RNA

and PNA the extinction coefficients in Table 7.1 were used. For PNA concentration determinations the sample was heated to 65 °C and UV measured to remove any secondary structure or association between strands. Macrocycle incorporations were attributed to having molar exctinction coefficients of 8,294 cm^{-1} and 10,700 cm^{-1} for the copper and nickel moieites respectively.[6] The ferrocene part of FcNA incorporations was considered to have an extinction coefficient of 3,500 cm^{-1}.[7]

Table 7.1 - Molar exctinction coeffcients of residues used to calculate oligonucleotide concentrations.

	Molar Extinction Coefficients $_{@\ 260\ nm}$ / $mol^{-1}\ cm^{-1}$	
Residue	**RNA**[3]	**PNA**[8]
Adenine	15400	13700
Guanine	13700	11700
Cytosine	9000	6600
Thymine	n/a	8600
Uridine	10000	n/a

7.3.2 Thermal Melting Experiments

All DNA thermal melts were measured on a Cary 5000 (Agilent Technologies Ltd, Cheshire, UK) UV/vis spectrophotometer fitted with a peltier temperature controller (temperature error is ± 0.2 °C). The temperature was ramped between 15 °C and 85 °C at a rate of 0.5 °C per minute and the absorbance changes measured at 260 nm unless otherwise stated. Each)experiment consisted of a minimum of two dissociations and one annealing ramp. The T_m was calculated from the average first derivative of the dissociation curves. For all samples the temperature observed is the sum of a minimum of 3 experimental repeats. Samples were prepared as 0.5 ml solutions as follows; 5 μM of each oligonucleotide strand, 100 mM $NaClO_4$ (FcNA) or 100 mM NaCl (Macrocycles) and 10 mM sodium phosphate buffer pH 7.0. For clarity and for purposes of experimental accuracy thermal melting temperatures are described to the nearest 0.5 °C (the error associated with the scan rate). In all cases the experimental error (standard error of the mean) is also reported in parenthesis.

Thermodynamic parameters, enthalpy change (ΔH^0), entropy change (ΔS^0) and Gibb's free energy (ΔG^0) calculated from thermal meltings curves were obtained using the hyperchromicity/van't Hoff method which has been previously described.[9,10]

7.3.3 Circular Dichroism Spectroscopy

Experiments were carried out on a JASCO J-810 spectropolarimeter (Tokyo, Japan). Measurements were made with a response time of 1.0 s, data interval of 0.2 nm, data pitch of 0.2 nm, bandwidth of 2 nm and a scanning speed of 200 nm/min with a total of 5 accumulations. Each data set was corrected for the baseline and noise. Samples were made up as follows 5 μM of each strand, 100 mM $NaClO_4$/NaCl, 10 mM sodium phosphate buffer pH 7.0 and were annealed at 85 °C and slowly cooled at room temperature before measurements.

7.3.4 Mercury binding

$Hg(ClO_4)_2$ was used throughout for all binding experiments studying the effects of thymine-Hg-thymine base pairing. $BaClO_4$, $PbClO_4$, $ZnNO_3$, $CdClO_4$, $CuSO_4$, $CaClO_4$, $MgSO_4$ and $Fe(ClO_4)_2$ were used as control metals. All metal solutions were made up fresh in ultra-pure water and used within a week to avoid changes in stock concentrations. Solutions were stored in ultra-clean glassware cleaned by soaking in basic piranha solution (CAUTION) for 2 hours before rinsing with copious amounts of ultra-pure water. Extra care should be taken when handling Cd, Pb and Hg salts.

7.4 Electrochemistry

7.4.1 Materials

Electrodes and diamond slurries (10 μm , 3 μm and 1 μm) were purchased from BioAnalytical Systems Inc. (BASi) (West Lafeyette, USA) Alumina powders (1 μm, 0.3 μm and 0.05 μm) were purchased from Buehler (Coventry, UK). Water purified with a Millipore Elix-Gradient A10 system (resistivity > 18 μΩ cm toc ≤ 5 ppb, Millipore, France) was used

throughout all experiments and cleaning procedures. All electrochemical measurements were carried out with a BASi EC epsilon potentiostat using a BASi C3 cell stand.

7.4.2 Cleaning and Preparation

Before use, all glassware was cleaned with an acidic piranha solution (CAUTION: piranha solution is an extremely strong oxidizing agent and should never be used alone) for 10 minutes, then rinsed with copious amounts of ultra-pure water and soaked again overnight in ultra-pure water. Cleaning with piranha is essential to the reproducibility of these experiments. The following day the glassware was then thoroughly rinsed again with ultra-pure water and placed in a clean oven to dry before use.

Platinum wire counter electrodes were cleaned before use using a Bunsen flame followed by rinsing with copious amounts of ultra-pure water. Ag/AgCl reference electrodes were stored between experiments in a 3 M KCl solution and rinsed thoroughly before use. Working electrodes were initially polished with three grades of diamond slurry (10 μm, 3 μm and 1 μm) on a polishing pad (BASInc, USA) for 5 minutes by moving in a figure of eight fashion, followed by polishing using three grades of alumina slurry on microcloth polishing pads (Buehler, UK) for 5 minutes each. Between each grade the electrode was rinsed thoroughly with a stream of ethanol (diamond) or ultra-pure water (alumina). The polished electrodes were then sonicated in ultra-pure water for 10 minutes to remove residual particles. Following sonication, the electrodes were again rinsed thoroughly with ultra-pure water before electrochemical cleaning using a traditional 3-electrode set up (including a Ag/AgCl reference electrode and platinum counter electrode) and the appropriate procedure as described below;

For cleaning gold working electrodes;

Potential cycles using cyclic voltammetry were carried out in a deoxygenated 0.5 M NaOH solution between -0.35 mV and -1.35 mV at a scan rate of 1000 mV s^{-1} until the

voltammograms were consistent (*ca.* 400 scans). The electrodes were then rinsed with ultra-pure water and cycled in 0.05 M H_2SO_4 solution between -0.35 V and 1.6 V with a scan rate of 100 mV s^{-1} until the voltammograms remained the same (*ca*. 16 scans). The surface roughness and hence geometric area were from integration of the gold oxide reduction peak measured from a CV in 0.05 M H_2SO_4 at 100 mV s^{-1} between -0.35 V and 1.5 V and using a literature value of 482 μC cm^{-2} to evaluate active electrochemical area.[11] The electrodes were then rinsed with ultra-pure water and dried under a stream of argon.

For glassy carbon working electrodes;

CVs were carried out in deoxygenated 0.05 M H_2SO_4 between 2 V and -2 V at 100 mV s^{-1} until the voltammograms remained consistent (*ca.* 50 scans). The electrodes were then rinsed with ultra-pure water and dried under a stream of argon.

Ferrocene Nucleic Acid Conjugates

Electrochemical measurements were carried out using a traditional 3-electrode set up utilising a polycrystalline gold working electrode (1.6 mm diameter), Ag/AgCl reference electrode and platinum wire counter electrode. The composition of samples was as follows: 50 μM oligonucleotide, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$ in ultra-pure water. Each sample was bubbled with argon for 30 minutes prior to measurements to remove oxygen.

Scan rate dependencies of FcNA materials was measured using cyclic voltammetry and scanning between appropriate potentials at scan rates between 5 and 50 mV s^{-1}. Upon addition of target sequences (titres up to 1.2 molar equivalents), 15 minutes of hybridization time was allowed after each addition to ensure the majority of target was bound. Square wave voltammetry (1 mV step, 25 mV amplitude and frequency of 8 Hz) was used to probe the sensing ability of FcNA systems.

Macrocycles

Measurements were carried out using a traditional 3-electrode set up with a glassy carbon working electrode (3.0 mm diameter), Ag/AgCl reference electrode and platinum wire counter electrode. The samples composition was as follows: 50 μM oligonucleotide, 10 mM sodium Tris-HCl buffer pH 7.0, 100 mM NaCl in ultra-pure water. Each sample was bubbled with argon for 30 minutes prior to measurements to remove oxygen.

Scan rate dependences were monitored using cyclic voltammetry between 1 and 20 mV s^{-1} (the capacitance/background current observed increased dramatically at rates above 50 mv s^{-1}). Between target additions (titres up to 1.2 molar equivalents), 15 minute hybridization time was allowed before measurements.

Data Analysis

All electrochemical data was analysed using OriginPro 9.1 (Massachussets, U.S.A). SWV voltammograms were background subtracted using the same software and the original data is shown in the appendices.

7.5 Molecular Dynamics

Molecular Dynamic simulations were carried out by Dr. Jean-Louis Duprey. In brief, models were undertaken using AMBER II modelling suite on the University of Birmingham Bluebear computer cluster. The structures were first energy minimised for (500 ps) and subsequently submitted for molecular dynamic optimisation for a period of no less than 5 ns using an explicit solvent system (the exact time is detailed for each model within the text). The pdb files from the generated models were then depicted using Chimera II graphical software.[12]

7.6 DNA Labelled Virus Particles

This experimental section is taken from a recently published article in *ACS Synthetic Biology*.[13]

7.6.1 Growth and Purification of Wild-type M13 – Carried out by Dr. Pacheco-Gomez

Wild-type M13 bacteriophage (wtM13) was produced by inoculating *E. coli* Top 10 F' using the method detailed previously.[14] To purify the M13 bacteriophage the material recovered from the first round of purification was subjected to a further purification using caesium chloride gradient centrifugation. M13 bacteriophage was mixed with CsCl powder to give a final concentration of 0.4 mg/mL of CsCl. The samples were centrifuged for 24 hours at 35,000 rpm in a 70.1 Ti rotor at 15 °C. This resulted in the formation of a diffuse band containing M13 bacteriophage. The band was extracted and the samples were dialysed in 50 mM potassium phosphate buffer pH 8.0 with 3 changes of 1 L each.

7.6.2 M13 Conjugation of SMCC.

An initial mass of 2 mg (1.1 x 10^{-10} moles) of CsCl purified wtM13 bacteriophage was adjusted to a final volume of 1 ml by adding 0.1 M potassium phosphate buffer, pH 7.5 in a micro-centrifuge tube. 1.056 mg (3.16 x 10^{-6} moles) of SMCC (Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate, Thermo Scientific Pierce) was dissolved in 10 µL DMSO and added to the mixture to provide a 10:1 molar excess over the M13 bacteriophage pVIII protein (it is assumed that there are 2700 pVIII proteins per M13 particle). The sample was then incubated for 1 hour at room temperature, before being quenched with 0.23 mg (3.16 x 10^{-6} moles) of glycine for 30 minutes at room temperature. The maleimide-derivatised product was then separated from low molecular mass by-products using a PD-10 column (Sephadex G-25, GE Healthcare). The sample was eluted in conjugation buffer (50 mM potassium phosphate buffer, 150 mM NaCl, 5 mM EDTA, pH 7.0).

7.6.3 Oligonucleotide Synthesis, Purification and Characterisation.

Synthesis of the rhodamine labeled forward primer (**R-For**) 5'-SSC_6-(dT-carboxy tetramethyl-rhodamine)-ATG AGT ATT CAA CAT TTC-3', complementary sequence (**F-Com**) 5'-(6-FAM)- GAA ATG TTG AAT ACT CAT-3' and complete mismatch (**F-Ran**) 5'-(6-FAM)-TCA TCA GTC AGT CAG TCA-3' was carried out on an Applied Biosystems 394 DNA/RNA synthesiser (Foster City, California). ULTRA-mild DNA synthesis conditions were used for the synthesis of fluorophore labeled oligonucleotides in which the use of Pac-dA-CE, Ac-dC-CE and iPr-Pac-dG-CE phosphoramadites was required. Cleavage from CPG column was carried out using 0.05 M potassium carbonate solution at RT for 4 hours. Labeling of oligonucleotides with a 5' disulfide (SSC_6) was achieved *via* a commercially available C6 trityl protected disulfide phosphoramidite (Glen Research, Virginia, USA). Fluorophore labeling with carboxy tetramethyl-rhodamine(TAMRA)-dt and 6-Carboxyfluorescein(FAM) was achieved via commercially available phosphoramadites (Glen Research). Purification and characterisation was carried out as previously described.[15] Disulfide labeled oligonucleotides were deprotected according to manufacturer's instructions and subsequently desalted for a second time using a NAP-10 column. The disulfide labeled oligonucleotides were then reacted with 50 mM tris(2-carboxyethyl)phosphine (TCEP) to reduce the disulfide to a reactive thiol for 30 minutes prior to biocojugation.

7.6.4 Bioconjugation.

11.7 µL of 1 mM (1.17×10^{-8} moles) thiol labeled oligonucleotide in conjugation buffer (50 mM potassium phosphate buffer, 150 mM NaCl, 5 mM EDTA, pH 7.0) was added to the SMCC labeled M13. The solution was mixed and left to react at RT for 1 hour before gentle stirring at 5 ^{o}C overnight. 14.6 µL of a 10 mg/ml (1.17×10^{-7} moles) solution of N-ethylmaleimide (NEM) in conjugation buffer was then added to the mixture, stirred and left to react at RT for 15 minutes and finally purified by size exclusion chromatography.

7.6.5 Size Exclusion Chromatography.

M13-oligonucleotide conjugates were purified using an ÄKTA Explorer 10 purification system (GE Healthcare), fitted with a Superdex 200 (120 mL column volume) HiLoad 16/60 prepacked column. The column was equilibrated with 1 column volume of water and 1.2 column volumes of buffer (50 mM potassium phosphate buffer and 150 mM NaCl, pH 8.0) before sample injection. Three wavelengths were recorded simultaneously: λ_1 = 269 nm, λ_2 = 280 nm and λ_3 = 556 nm. 2 mL fractions were collected using a Frac-950 fraction collector (Amersham Pharmacia Biotech).

7.6.6 SDS PAGE - Carried out by Haydn Little.

Samples were run vertically on a 12% SDS Page denaturing gel. A 20 µL sample (0.8 mg/ml M13) in loading buffer (Tris, glycerol, β-mercaptoethanol) was first denatured at 95 ^{o}C for 15 minutes before loading into the wells. The gel was run for 40 minutes at 200 V (400 mA) before staining with Coomassie Blue for 1 hour, and subsequently destained using a 3x solution of acetic acid in MeOH for 1 hour. The resulting gel was then visualised on a light box.

7.6.7 Measurement of M13/DNA Concentration.

Absorption spectra to determine bacteriophage concentrations for LD experiments using the method previously detailed, were carried out using a JASCO V550 UV-Vis spectrometer. Conjugation efficiencies were calculated using the observed rhodamine absorbance at 556 nm (89,000 M^{-1} cm^{-1}).[3]

7.6.8 Measurement of LD Spectra – Carried out in collaboration with Nadja Steinke

Measurements were performed on a Jasco J-715 Spectropolarimeter (Jasco, Japan) which has been modified to measure linear dichroism. The samples were placed into a micro-Couette cell housed in the Jasco J-715. LD spectra were recorded at room temperature over a wavelength range from 350−190 nm at a scan speed of 100 nm/min, with a 0.2 nm data pitch, 2.0 nm band width and 1 sec response time. Couette flow was generated through a

rotation at 3000 rpm. For each experiment, a baseline (non-rotating capillary) was collected, subtracted from the rotating spectrum and the signal was subsequently zeroed at 350 nm.

7.6.9 Polymerase Chain Reaction (PCR) – Carried out in collaboration with Dr. Pacheco-Gomez

All PCR reactions were carried out using a Biometra(R) T 3000 Thermocycler. 100 ng of pcDNA template was used in a final reaction volume of 50 μL. Reaction mixture included either the unconjugated **R-For** oligonucleotide or **R-For** conjugated to M13 bacteriophage and antisense oligonucleotide primers (2 pmol/μl); 200 μM dNTPs, 0.2 nmol/ul; GoTaq DNA polymerase (5 units) in 1x GoTaq Flexi Green buffer. 1.5 μL of ethylene glycol was added to allow for amplification at lower temperatures.[16] DNA was amplified using the following conditions: 86.0 °C for 5 min (preheating); denaturing 86.0 °C, 1 min; annealing 50.0 °C, extension 72.0 °C, 1 min. Repeat from denaturation 35 times. Followed by a final extension 72.0 °C for 5 minutes. Lid temperature was set at 86.0 °C.

7.6.10 Agarose Gel Electrophoresis - Carried out in collaboration with Dr. Pacheco-Gomez

The PCR mixtures were run on a horizontal 1% (w/v) agarose gel using 2 x TAE buffer. Ethidium bromide was added to the gel before the gel had set to visualize the PCR product(s).10 μL of each sample was loaded onto the gel in 2 x TAE buffer. A 1 kb ladder DNA ladder was used as a reference. The gel was subsequently run at 80 V for 30 minutes, before being visualized on a light box.

7.7 References

1. I. Mames, U. E. Wawrzyniak, M. Woźny, R. Bilewicz and B. Korybut-Daszkiewicz, *Dalton Trans.*, 2013, **42**, 2382-2391.
2. A. Więckowska, M. Wiśniewska, M. Chrzanowski, J. Kowalski, B. Korybut-Daszkiewicz and R. Bilewicz, *Pure Appl. Chem.*, 2007, **79**, 1077-1085.
3. W. A. Kibbe, *Nucleic Acids Res.*, 2007, **35**, W43-W46.
4. C. R. Cantor, M. M. Warshaw and H. Shapiro, *Biopolymers*, 1970, **9**, 1059-1077.
5. M. J. Cavaluzzi and P. N. Borer, *Nucleic Acids Res.*, 2004, **32**, e13-e13.
6. J. Kowalski, Editon edn.
7. Z.-y. Zhao, University of Birmingham, Editon edn., 2009.
8. L. Christensen, R. Fitzpatrick, B. Gildea, K. H. Petersen, H. F. Hansen, T. Koch, M. Egholm, O. Buchardt, P. E. Nielsen, J. Coull and R. H. Berg, *J. Pept. Sci.*, 1995, **1**, 175-183.
9. J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.
10. A. Sen and P. E. Nielsen, *Nucleic Acids Res.*, 2007, **35**, 3367-3374.
11. J. C. Hoogvliet, M. Dijksma, B. Kamp and W. P. van Bennekom, *Anal. Chem.*, 2000, **72**, 2016-2021.
12. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
13. J. Carr-Smith, R. Pacheco-Gomez, H. A. Little, M. R. Hicks, S. Sandhu, N. Steinke, D. J. Smith, A. Rodger, S. A. Goodchild, R. A. Lukaszewski, J. H. R. Tucker and T. R. Dafforn, *ACS Syn. Bio.*, 2015, 10.1021/acssynbio.5b00034, In Press.
14. R. Pacheco-Gómez, J. Kraemer, S. Stokoe, H. J. England, C. W. Penn, E. Stanley, A. Rodger, J. Ward, M. R. Hicks and T. R. Dafforn, *Anal. Chem.*, 2012, **84**, 91-97.
15. H. V. Nguyen, Z.-y. Zhao, A. Sallustrau, S. L. Horswell, L. Male, A. Mulas and J. H. R. Tucker, *Chem. Commun.*, 2012, **48**, 12165-12167.
16. A. M. Schwartz and G. D. Fasman, *Biopolymers*, 1979, **18**, 1045-1063.

Appendices

8.1 Synthesised Oligonucleotides for FcNA work

Appendix 8-1 - Synthesised DNA Oligonucleotides.

S1 = 5'-TGGACTC**XX**CTCAATG-3'					
Code	**XX =**	**Predicted Mass / Da**	**Observed Mass / Da**	**HPLC Method**	**Retention time / min**
S1TT	TT	4847	4847	Oligo60	32.46
S1AA	AA	4865	4865	Oligo60	24.99
S1GG	GG	4897	4898	ThiolFerro	17.70
S1CC	CC	4817	4818	ThiolFerro	17.78
S1AbAb	AbAb	4599	4599	Oligo60	25.74
S1FcHH$_{RR}$	**FcHH$_{RR}$**	4631	4631	Oligo60	30.09
Code	**XX =**	**Predicted Mass / Da**	**Observed Mass / Da**	**HPLC Method**	**Retention time / min**
S1FcHH$_{SS}$	**FcHH$_{SS}$**	4631	4631	Oligo69	30.15
S1Fcn3c2TT$_{RRpSpS}$ Also referred to as S1FcTT$_{RR}$	**Fcn3c2TT$_{RRpSpS}$**	4936	4936	ThiolFerro	29.36
S1Fcn2c3TT$_{SSpRpR}$ also referred to as S1FcTT$_{SS}$	**Fcn3c2TT$_{SSpRpR}$**	4936	4936	ThiolFerro	29.97
S1Fcn3c3TT$_{RRpSpS}$	**Fcn3c3TT$_{RRpSpS}$**	4961	Oligo 993	ThiolFerro	18.08
S1Fcn3c3TT$_{SSpRpR}$	**Fcn3c3TT$_{SSpRpR}$**	4961	Oligo 703	ThiolFerro	18.03
S1Fcn3c2TT$_{pRpR}$	**Fcn3c2TT$_{pRpR}$**	4905	Oligo 942	ThiolFerro	15.93
S1Fcn3c2TT$_{pSpS}$	**Fcn3c2TT$_{pSpS}$**	4905	Oligo 996	ThiolFerro	16.87
S1Fcn2c2TT$_{pRpR}$	**Fcn2c2TT$_{pRpR}$**	4879	4879	Oligo60	35.47
S1FcCC$_{RR}$	**FcCC$_{RR}$**	4955	4955	ThiolFerro	27.21
S1FcAA$_{RR}$	**FcAA$_{RR}$**	4954	4953	Oligo60	38.52

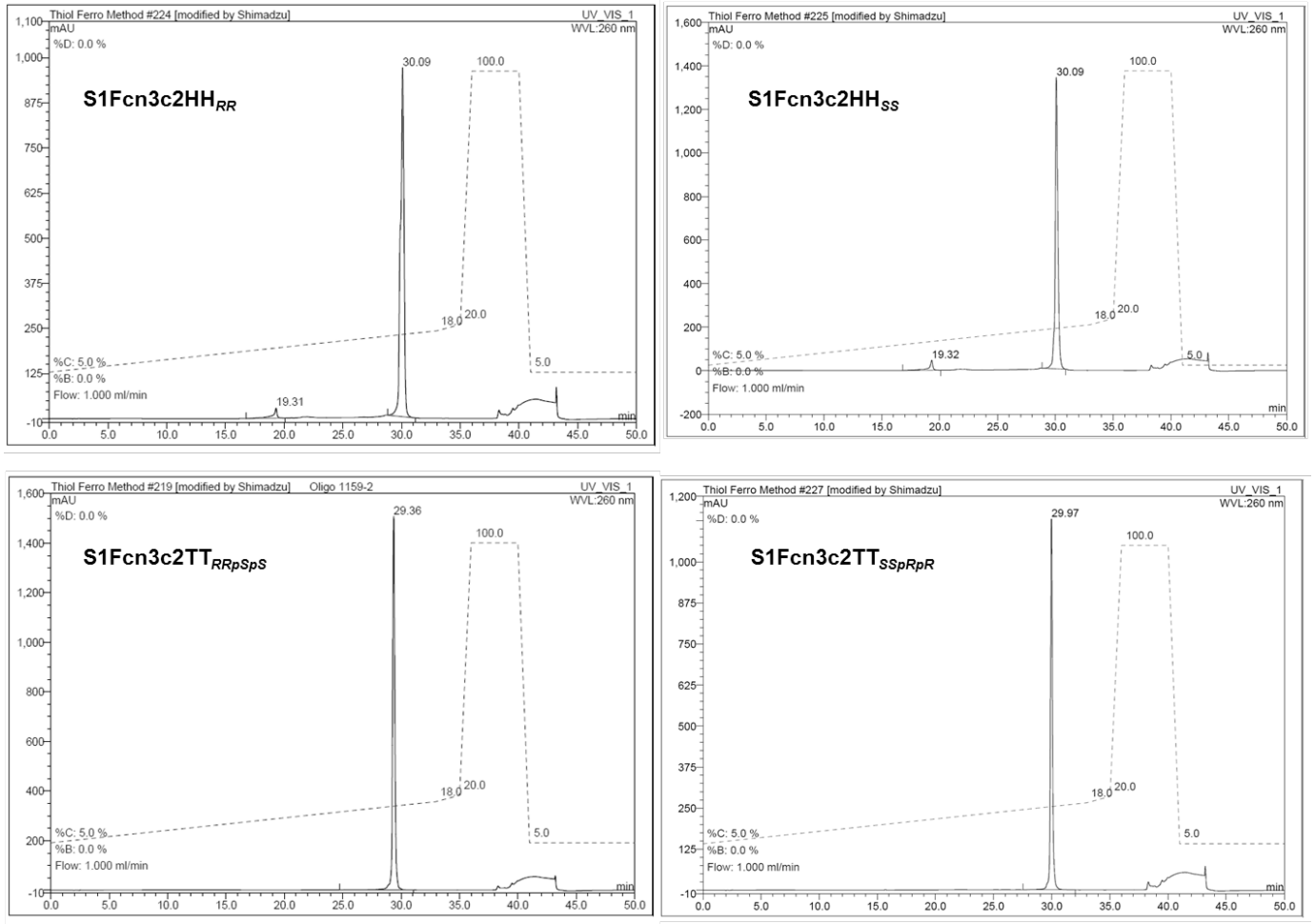
S1FcAA$_{SS}$	**FcAA$_{SS}$**	4954	4953	Oligo60	39.09
S2 = 5'-CATTGAG**YY**GAGTCCA-3'					
Code	**YY =**	**Predicted Mass / Da**	**Observed Mass / Da**	**HPLC Method**	**Retention time / min**
S2TT	TT	4896	4897	Oligo60	31.35
S2AA	AA	4914	4916	Oligo60	22.85
S2GG	GG	4946	4946	Oligo60	28.05
S2CC	CC	4866	4866	ThiolFerro	18.57
S2AG	AG	4930	4930	Oligo60	26.79
S2GA	GA	4930	4930	Oligo60	27.31
S2AbAb	AbAb	9649	9649	Oligo60	29.13
S2UUR (RNA)	UU	5082	5083	ThiolFerro	15.09
S2AAR (RNA)	AA	5129	5129	ThiolFerro	15.45
S2GGR (RNA)	GG	5161	5129	ThiolFerro	10.29
S2CCR (RNA)	CC	5081	5080	ThiolFerro	10.35
S2TTP (PNA)*	TT	4496	4498	Purified by Anjana Sen University of Copenhagen.**	
S2AAP (PNA)*	AA	4514	4515		
S2GGP (PNA)*	GG	4546	4547		
S2AGP (PNA)*	AG	4530	4530		
S2FcTT$_{RR}$	**FcTT$_{RR}$**	4985	4985	Oligo60	53.38
S2FcTT$_{SS}$	**FcTT$_{SS}$**	4985	4985	Oligo60	51.21
S2FcAA$_{RR}$	**FcAA$_{RR}$**	5002	5003	ThiolFerro	16.75
S2FcAA$_{SS}$	**FcAA$_{SS}$**	5002	5003	ThiolFerro	15.90

**Synthesised by Dr. Anjana Sen – University of Copenhagen.

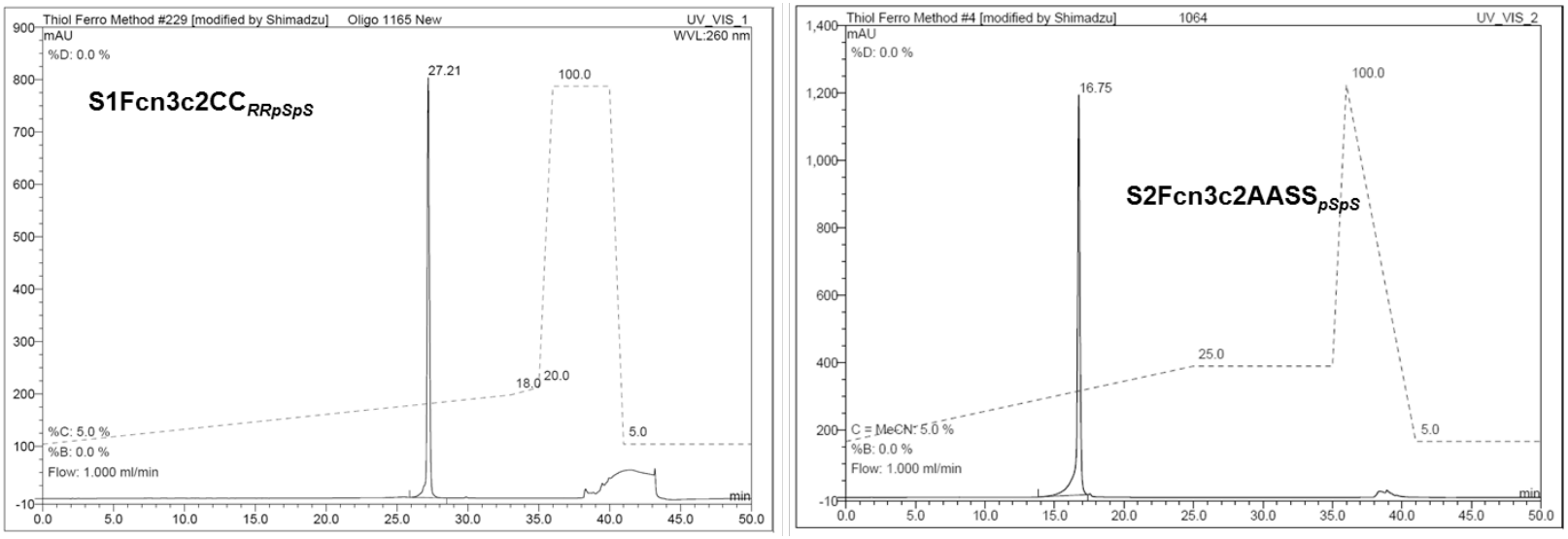
Included below are some example HPLC traces of the above described sequences.

HPLC methods used the following gradients;

- ThiolFerro: A two solvent system using Buffer D and MeCN gradients was used in the following time frame;
    - 5-20% MeCN: 35 minutes, 100% MeCN: 7 minutes, 5% MeCN: 8 minutes. Solvent rate: 3 ml per minute.
- Oligo60: A three solvent system using Buffer A, Buffer D and MeCN gradients was used in the following time frames;
    - 30-50% Buffer B: 25 minutes, 50-100% Buffer B: 7 minutes, 100-0% Buffer B: 13 minutes, 0% Buffer B: 10 minutes, 30% Buffer B, 5 minutes
    - 0% MeCN: 35 minutes, 0-100% MeCN: 10 minutes, 100% MeCN: 10 minutes, 0% MeCN: 5 minutes.



Appendix 8-2 - HPLC traces of both stereoisomers of ***S1Fcn3c2HH$_{XX}$*** *and* ***S1Fcn3c2TT$_{XXpYpY}$****.*



Appendix 8-3 - HPLC traces of ***S1Fcn3c2CC$_{RRpSpS}$*** *and* ***S2Fcn3c2AA$_{SSpRpR}$***

.

8.2 Ferrocene Nucleic Acids: A DNA Mimic

Appendix 8-4 - All of the T_m values obtained for FcNA-DNA conjugates with DNA target S2YY. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.

	T_m / °C								
S1XX =	**S2TT**	**S2AA**	**S2GG**	**S2CC**	**S2AbAb**	**S2AG**	**Opposite Itself**	**$FcAA_{SS}$**	**$FcAA_{RR}$**
$FcHH_{RR}$[$$]	36.5	37.5	36.5	36.0	35.5	36.5	42.0	-	-
$FcHH_{SS}$	36.0	37.5	36.5	35.5	37.0	36.5	43.0	-	-
$FcTT_{RR}$	40.5	43.0	41.0	-	38.5	42.0	44.5	46.5	46.0
$FcTT_{SS}$	39.5	41.0	39.5	-	37.0	39.5	45.5	49.5	46.0
$FcCC_{RR}$	37.5	39.5	48.0	-	37.0	45.5	-	-	44.5
$FcAA_{RR}$	44.5	42.0	42.5	43.0	-	-	-	-	45.0
$FcAA_{SS}$	47.0	41.0	41.5	38.5	39.5	-	45.5	45.5	45.0
AbAb	34.0	37.0	37.0	-	-	37.0	-	-	-
TT	41.5	56.0	44.5	-	33.0	48.5	41.5	47.5	-
CC	38.0	38.0	60.0	36.0	32.5	50.5	36.0		
AA	57.5	-	48.5	38.5	37.5	-	-	42.0	-
Fcn2HH	35.5	36.5	37.0	34.5	-	-	-	-	-
$Fcn3c3TT_{RRpSpS}$	40.5	41.5	40.5	-	45.0	41.0	-	48.0	-
$Fcn3c3TT_{SSpRpR}$	40.5	43.0	40.5	-	41.5	41.5	47.0	49.0	-
$Fcn3c2TT_{pRpR}$	40.5	41.5	41.0	-	41.5	41.5	43.0	47.0	-
$Fcn3c2TT_{pSpS}$	40.5	40.5	40.0	-	42.0	39.5	47.0	46.5	-
$Fcn2c2TT_{pRpR}$	41.0	42.0	40.5	-	-	-	45.5	-	-

*Highlighted values represent the fully complementary sequences.

*Appendix 8-5 - All of the T_m values obtained for FcNA-DNA conjugates with RNA targets **S2YY**. 5 µM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C* **S2YYR**			
S1XX =	**S2UUR**	**S2AAR**	**S2GGR**	**S2CCR**
$FcHH_{RR}$[$$]	35.0	39.0	36.0	34.5
$FcHH_{SS}$	35.0	39.5	36.0	34.5
$FcTT_{RR}$	40.0	47.5	43.0	-
$FcTT_{SS}$	41.5	47.5	42.5	-
$FcCC_{RR}$	39.5	42.5	49.5	-
$FcAA_{RR}$	44.0	43.0	-	40.5
$FcAA_{SS}$	50.3	43.0	-	43.0
AbAb	33.5	40.0	38.0	-
TT	44.5	59.0	52.5	39.5
CC	39.0	45.0	69.5	43.5
AA	55.0	42.5	46.0	42.5

*Highlighted values represent the fully complementary sequences.

*Appendix 8-6 – All of the T_m values obtained for FcNA-DNA conjugates with PNA target **S2YY**. 5 µM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	T_m / °C **S2YYP**				
S1XX =	**S2TTP**	**S2AAP**	**S2GGP**	**S2AGP**	**S2ATP**
$FcHH_{RR}$[$$]	55.0	60.0	59.0	59.0	57.0
$FcHH_{SS}$	-	-	-	-	-
$FcTT_{RR}$	61.0	68.5	62.0	64.5	65.0
$FcTT_{SS}$	59.0	67.5	62.0	65.0	66.0
$FcCC_{RR}$	57.0	61.5	71.0	64.0	59.0
$FcAA_{RR}$	61.0	61.5	61.5	61.0	60.0
$FcAA_{SS}$	69.5	59.0	59.0	-	67.5
AbAb	-	-	-	-	-
TT	69.0	81.0	64.5	62.0	69.0
CC	54.0	58.5	86.0	69.5	57.0
AA	74.0	64.5	64.5	62.0	67.0

*Highlighted values represent the fully complementary sequences.

*Appendix 8-7 - Thermodynamic parameters for **S1XX•S2AA** (where **XX** = FcNA unit bearing thymines) calculated by van't Hoff analysis of the thermal melting profiles. 5 μM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM NaClO4.*

	S2AA			
S1 XX =	ΔH / kJmol^{-1}	ΔS / kJmol^{-1}	$\Delta G_{20°C}$ / kJmol^{-1}	T_{m**} / °C
FcHH$_{RR}$ $^{\$\$}$	-323.1 (±20.2)	-0.935 (±0.01)	-49.1 (±2.2)	38.7 (±0.5)
FcHH$_{SS}$ $^{\$\$}$	-362.2 (±24.9)	-1.058 (±0.08)	-52.2 (±0.7)	39.4 (±0.8)
FcTT$_{RR}$	-394.1 (±10.9)	-1.139 (±0.04)	-60.2 (±0.9)	44.6 (±1.1)
FcTT$_{SS}$	-272.6 (±17.3)	-0.764 (±0.05)	-47.8 (±1.4)	40.8 (±0.5)
Fcn3c3TT$_{RRpSpS}$ $^{\#}$	-406.0 (±15.0)	-1.179 (±0.04)	-60.7 (±2.1)	44.2 (±1.0)
Fcn3c3TT$_{SSpRpR}$ $^{\#}$	-428.2 (±38.2)	-1.245 (±0.12)	-63.5 (±3.7)	45.1 (±1.5)
Fcn3c2TT$_{pRpR}$	-402.2 (±14.5)	-1.171 (±0.05)	-59.1 (±1.2)	43.1 (±0.4)
Fcn3c2TT$_{pSpS}$	-386.9 (±6.7)	-1.129 (±0.02)	-56.2 (±1.1)	41.5 (±0.7)
TT	-475.1 (±17.0)	-1.340 (±0.05)	-82.4 (±1.9)	56.5 (±0.1)

$

Lower baseline determination not accurate. ** *T*m calculated from the thermodynamic parameters not the first derivative of the melting curve. $^{\#}$ Unstable FcNA monomer. Temperature increased with each ramp.

*Appendix 8-8 - Thermodynamic parameters for **S1XX•S2YY** (where **YY** = full complement /double mismatch) calculated by van't Hoff analysis of the thermal melting profiles. 5 µM of each strand, 10 mM sodium phosphate buffer pH 7.0, 100 mM $NaClO_4$.*

	S2AA				S2GG			
S1 XX =	ΔH / kJmol^{-1}	ΔS / kJmol^{-1}	ΔG$_{20°C}$ / kJmol^{-1}	T_{m**} / °C	ΔH / kJmol^{-1}	ΔS / kJmol^{-1}	ΔG$_{20°C}$ / kJmol^{-1}	T_{m**} / °C
FcHH$_{RR}$ $^{\$\$}$	-323.1 (±20.2)	-0.935 (±0.01)	-49.1 (±2.2)	38.7 (±0.5)	-315.5 (±1.0)	-0.909 (±0.02)	-49.2 (±2.2)	39.3 (±0.4)
FcTT$_{RR}$	-394.1 (±10.9)	-1.139 (±0.04)	-60.2 (±0.9)	44.6 (±1.1)	-407.3 (±18.5)	-1.192 (±0.06)	-58.1 (±0.8)	41.9 (±0.4)
FcCC$_{RR}$	-456.9 (±6.9)	-1.312 (±0.02)	-61.1 (±2.6)	44.3 (±2.2)	-456.9 (±6.9)	-1.312 (±0.02)	-72.5 (±1.0)	50.2 (±0.3)
AbAb	-446.4 (±6.9)	-1.123 (±0.02)	-51.6 (±0.2)	37.7 (±0.2)	-446.4 (±15.3)	-1.193 (±0.04)	-54.0 (±0.4)	39.0 (±0.2)
TT	-475.1 (±17.0)	-1.340 (±0.05)	-82.4 (±1.9)	56.5 (±0.1)	-418.5 (±1.6)	-1.215 (±0.01)	-65.1 (±3.7)	47.0 (±3.0)
CC	-385.9 (±3.3)	-1.135 (±0.01)	-53.3 (±0.3)	39.1 (±0.1)	-484.4 (±18.6)	-1.351 (±0.05)	-88.6 (±2.4)	60.6 (±0.1)

$$ - Lower baseline determination not accurate. ** Tm calculated from the thermodynamic parameters not the first derivative of the melting curve.

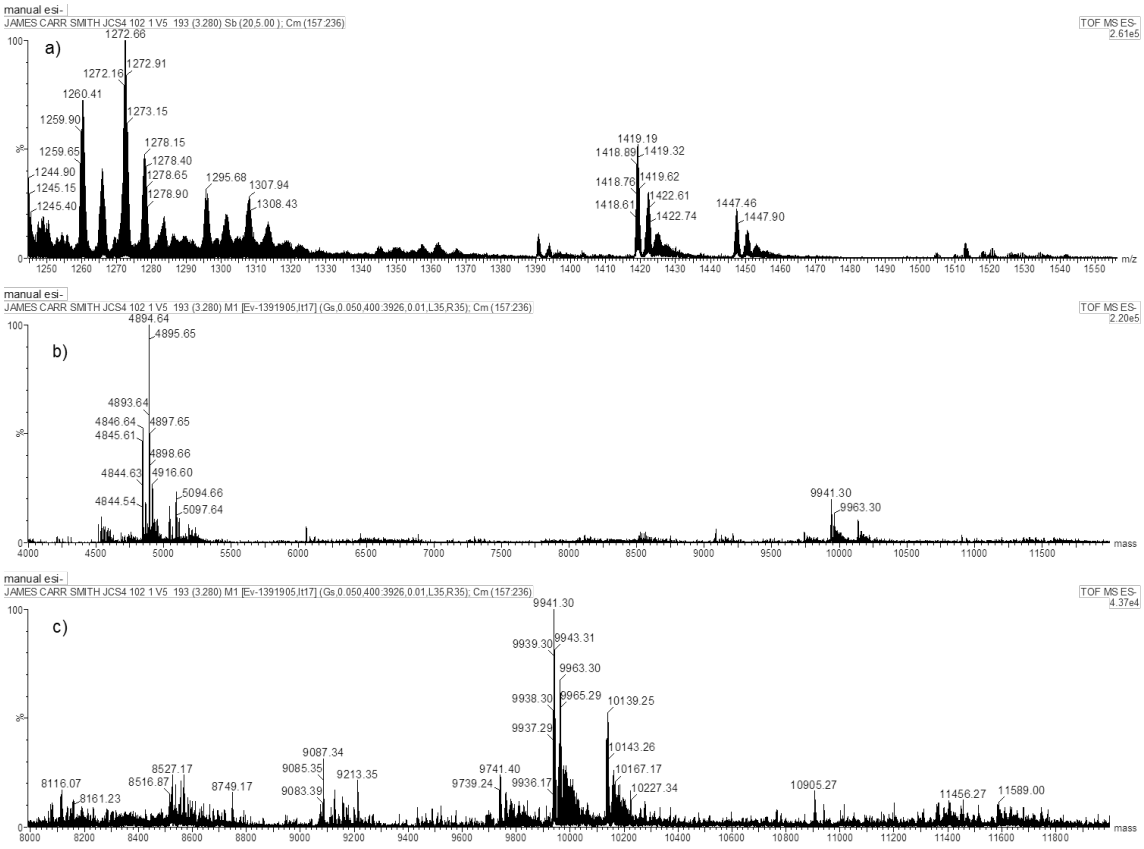
	S2TT				S2CC			
S1 XX =	ΔH / kJmol^{-1}	ΔS / kJmol^{-1}	ΔG$_{20°C}$ / kJmol^{-1}	T_{m**} / °C	ΔH / kJmol^{-1}	ΔS / kJmol^{-1}	ΔG$_{20°C}$ / kJmol^{-1}	T_{m**} / °C
FcHH$_{SS}$ $^{\$\$}$	-362.2 (±24.9)	-1.058 (±0.08)	-52.2 (±0.7)	39.4 (±0.8)	-397.0 (±18.5)	-1.181 (±0.05)	-51.0 (±1.6)	36.6 (±0.5
AbAb$^{\$\$}$	-385.9 (±1.5)	-1.153 (±0.01)	-48.1 (±0.1)	34.7 (±0.1)	-	-	-	-
FcAA$_{SS}$	-461.7 (±11.1)	-1.341 (±0.03)	-68.8 (±1.7)	47.1 (±0.7)	-422.2 (±4.2)	-1.247 (±0.01)	-56.9 (±0.2)	40.1 (±0.3)
AA	-478.9 (±3.2)	-1.347 (±0.01)	-84.1 (±1.1)	57.5 (±0.6)	-400.8 (±4.2)	-1.181 (±0.01)	-54.7 (±0.4)	39.5 (±0.1)

$$ - Lower baseline determination not accurate. ** Tm calculated from the thermodynamic parameters not the first derivative of the melting cur

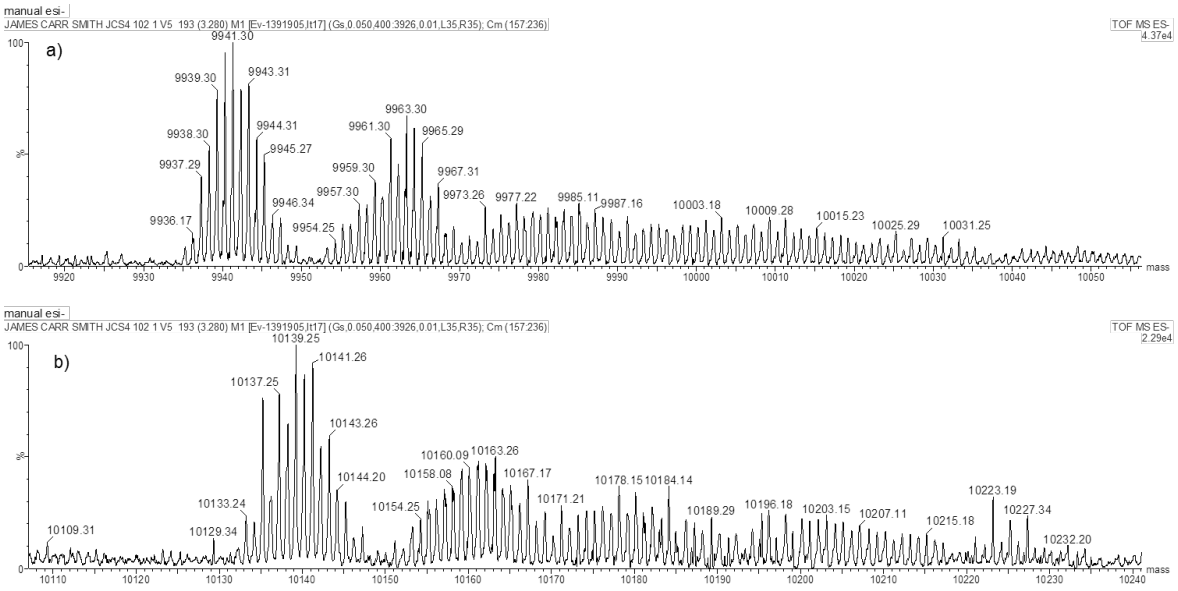
8.3 Metal Binding Ferrocene Nucleic Acids

Appendix 8-9 - The observed ions in the ESI mass spectrum when duplexes S1TTS2TT and S1FcTTS2TT are in the presence of mercury.

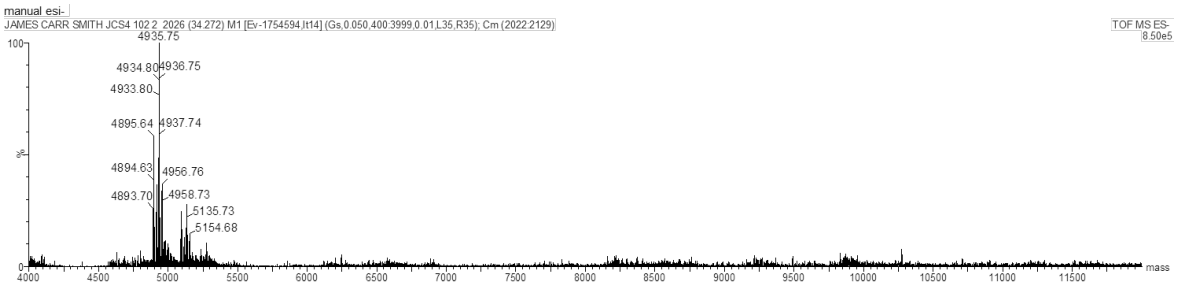
	Observed Mass / Da (Predicted Mass)		
	Duplex	**+ 1 T-Hg-T**	**+ 2 T-Hg-T**
S1TT•S2TT	9741 (9743)	9941 (9941)	10139 (10139)
S1FcTT$_{SS}$•S2TT	9832 (9832)	10030 (10030)	10228 (10228)



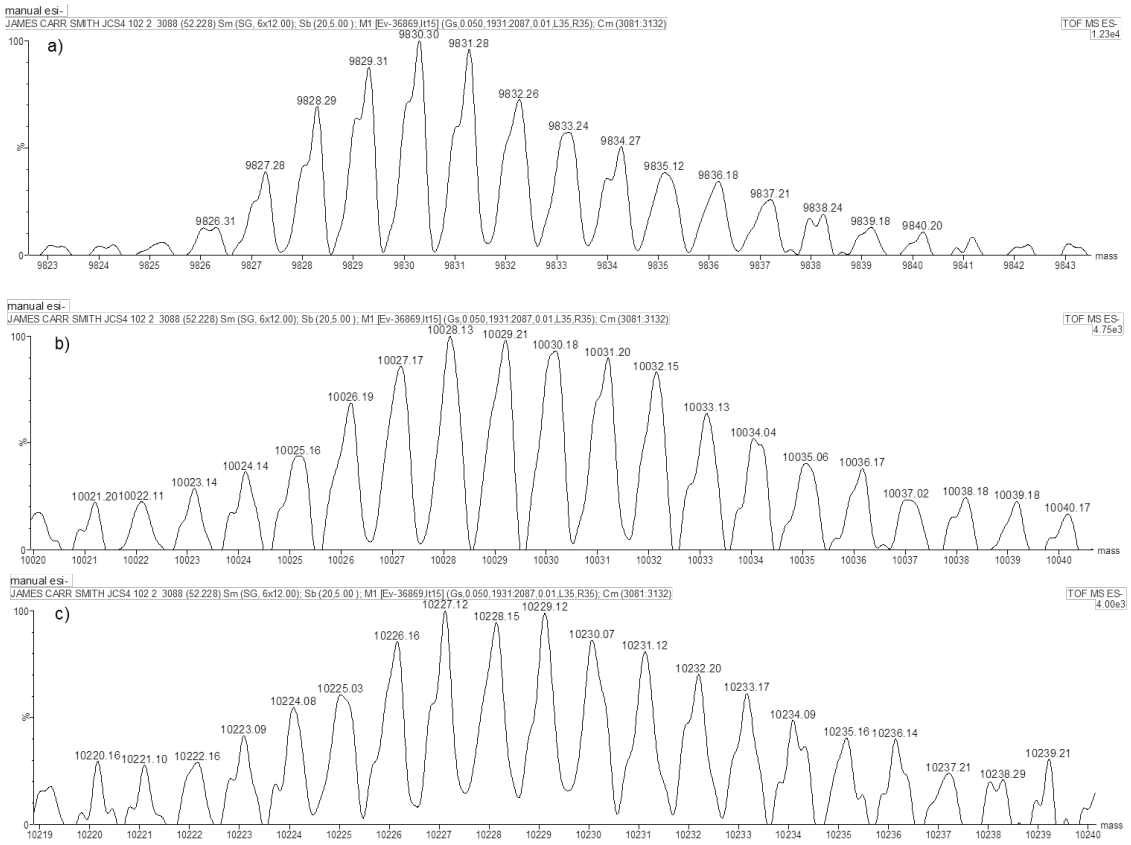
Appendix 8-10 - a) Raw data b) deconvoluted mass spectra and c) zoomed in region of the deconvoluted spectra for duplex S1TT-S2TT in the presence of 2 molar equivalents Hg2+, showing the ions of interest.



*Appendix 8-11 - Zoomed in regions of the deconvoluted mass spectra of duplex **S1TT•S2TT** in the presence of mercury. a) duplex only b) duplex – 2H + 1Hg c) duplex – 4H + 2Hg.*



*Appendix 8-12 - Raw data of the mass spectra obtained for duplex **S1FcTT$_{SS}$-S2TT** in the presence of mercury.*



Appendix 8-13 - Zoomed in regions of the deconvoluted mass spectra of duplex S1FcTTSS-S2TT in the presence of mercury showing the ions of interest. a) duplex only b) duplex – 2H + 1Hg c) duplex – 4H + 2Hg.

8.4 Redox Active Metal Centred Macrocycles

8.4.1 Oligonucleotide Synthesis

Appendix 8-14 – Synthesised Oligonucleotides for Chapter 5.

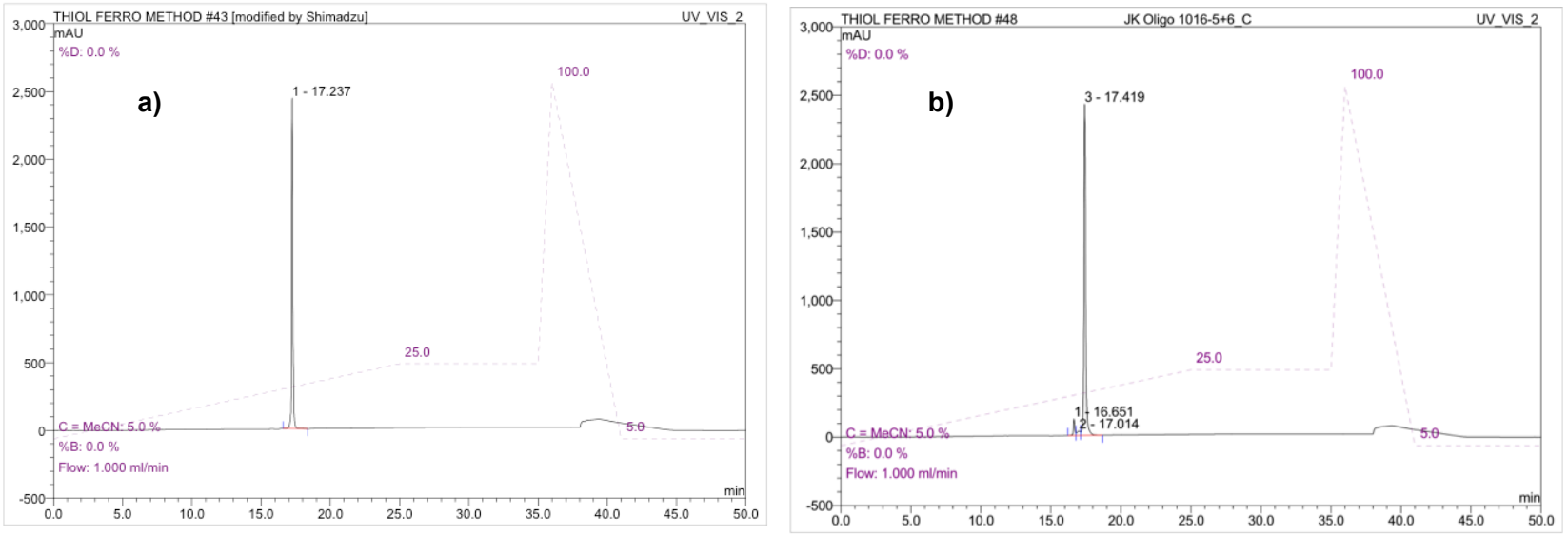
S1 = 5'-TGGACTC**X**CTCAATG-3'					
Code	**X =**	**Predicted Mass / Da**	**Observed Mass / Da**	**HPLC Method**	**Retention time / min**
S1Cu*a*	**Cu*a***	4771	4773	ThiolFerro	17.24
S1Cu*b*	**Cu*b***	4771	4773	ThiolFerro	17.42
S1Cu*a*Cu*a*	**2 x Cu*a***	5280	5278	ThiolFerro	19.72
S1Cu*b*Cu*b*	**2 x Cu*b***	5280	5278	ThiolFerro	20.71
S1Ni*a*	**Ni*a***	4767	4768	ThiolFerro	17.22
S1Ni*b*	**Ni*b***	4767	4768	ThiolFerro	17.49
S1Cu*L*	**Cu*L***	4759	4760	ThiolFerro	17.13
S1Ni*L*	**Ni*L***	4755	4754	ThiolFerro	16.78
S2 = 5'-CATTGAG**Y**GAGTCCA-3'					
Code	**Y =**	**Predicted Mass / Da**	**Observed Mass / Da**	**HPLC Method**	**Retention time / min**
S2T	T	4592	4592	Oligo60	17.78
S2A	A	4601	4601	Oligo60	16.82
S2G	G	4617	4619	Oligo60	16.80
S2C	C	4577	4577	Oligo60	17.22
S2Cu*a*	**Cu*a***	4822	4822	ThiolFerro	13.00
S2Cu*b*	**Cu*b***	4822	4822	ThiolFerro	18.34
S2Cu*a*Cu*a*	2 x **Cu*a***	5354	5355	ThiolFerro	18.31
S2Cu*b*Cu*b*	2 x **Cu*b***	5354	5355	ThiolFerro	21.31
S2Cu*L*Cu*L*	2 x **Cu*L***	5326	5327	ThiolFerro	18.90

S2Ni*L*Ni*L*	2 x **Cu*L***	5319	5319	ThiolFerro	19.45

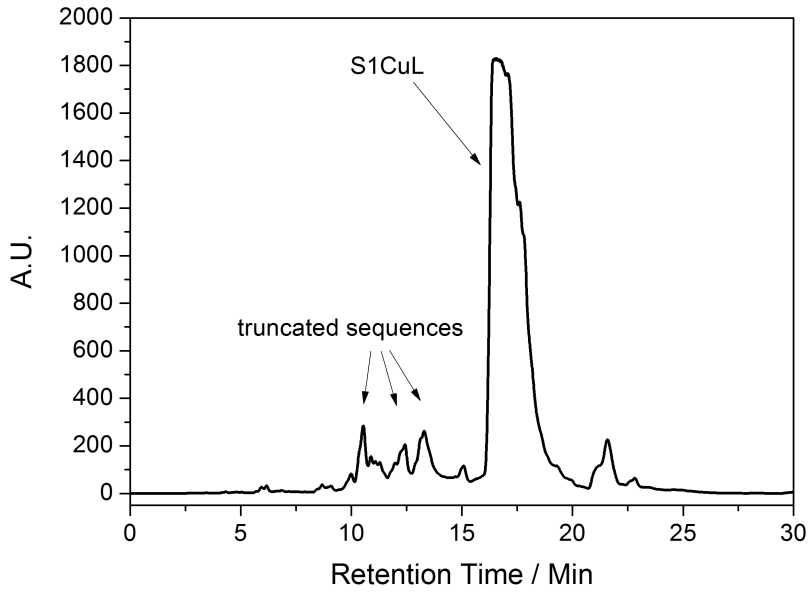
- *** Macrocycle modified sequences were purified and characterized in collaboration with Dr. Jaroslaw Kowalski and Dr. Jean-LouisDuprey

HPLC methods used the following gradients;

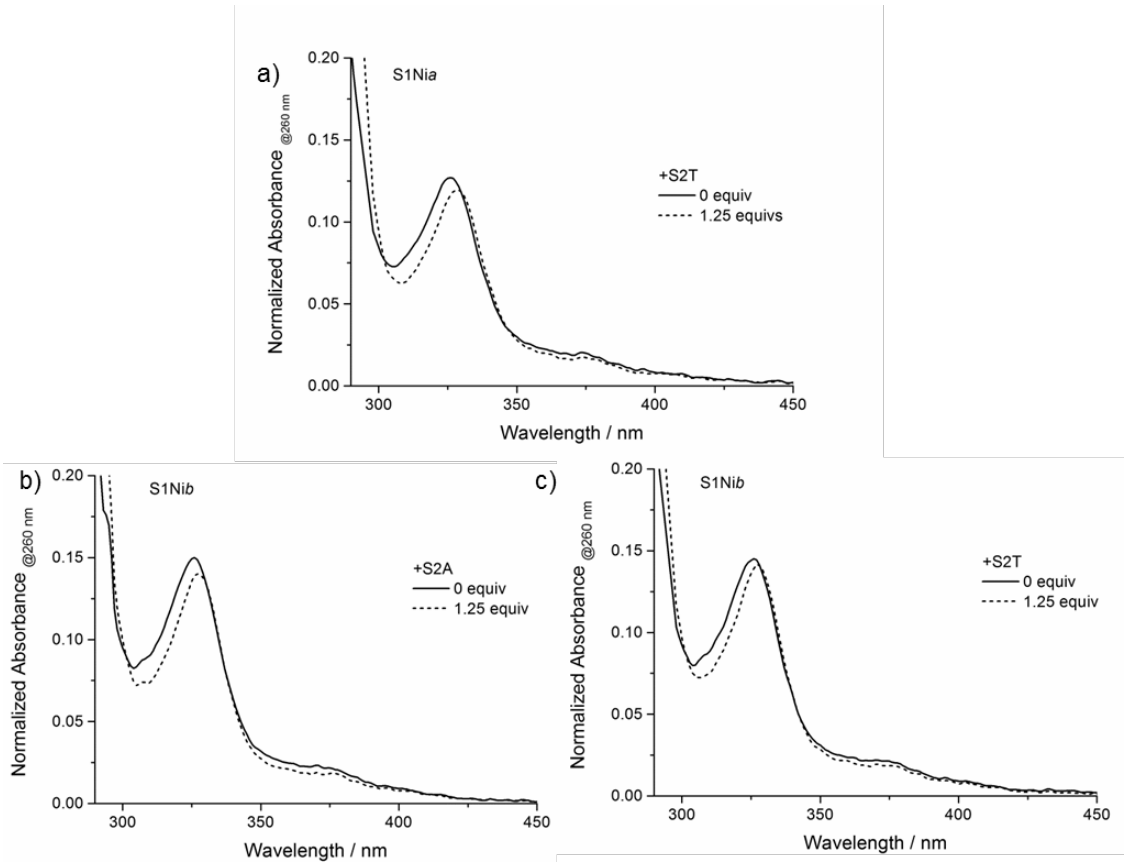
- ThiolFerro: A two solvent system using Buffer D and MeCN gradients was used in the following time frame;
    - 5-20% MeCN: 35 minutes, 100% MeCN: 7 minutes, 5% MeCN: 8 minutes. Solvent rate: 3 ml per minute.
- Oligo60: A three solvent system using Buffer A, Buffer D and MeCN gradients was used in the following time frames;
    - 30-50% Buffer B: 25 minutes, 50-100% Buffer B: 7 minutes, 100-0% Buffer B: 13 minutes, 0% Buffer B: 10 minutes, 30% Buffer B, 5 minutes
    - 0% MeCN: 35 minutes, 0-100% MeCN: 10 minutes, 100% MeCN: 10 minutes, 0% MeCN: 5 minutes.



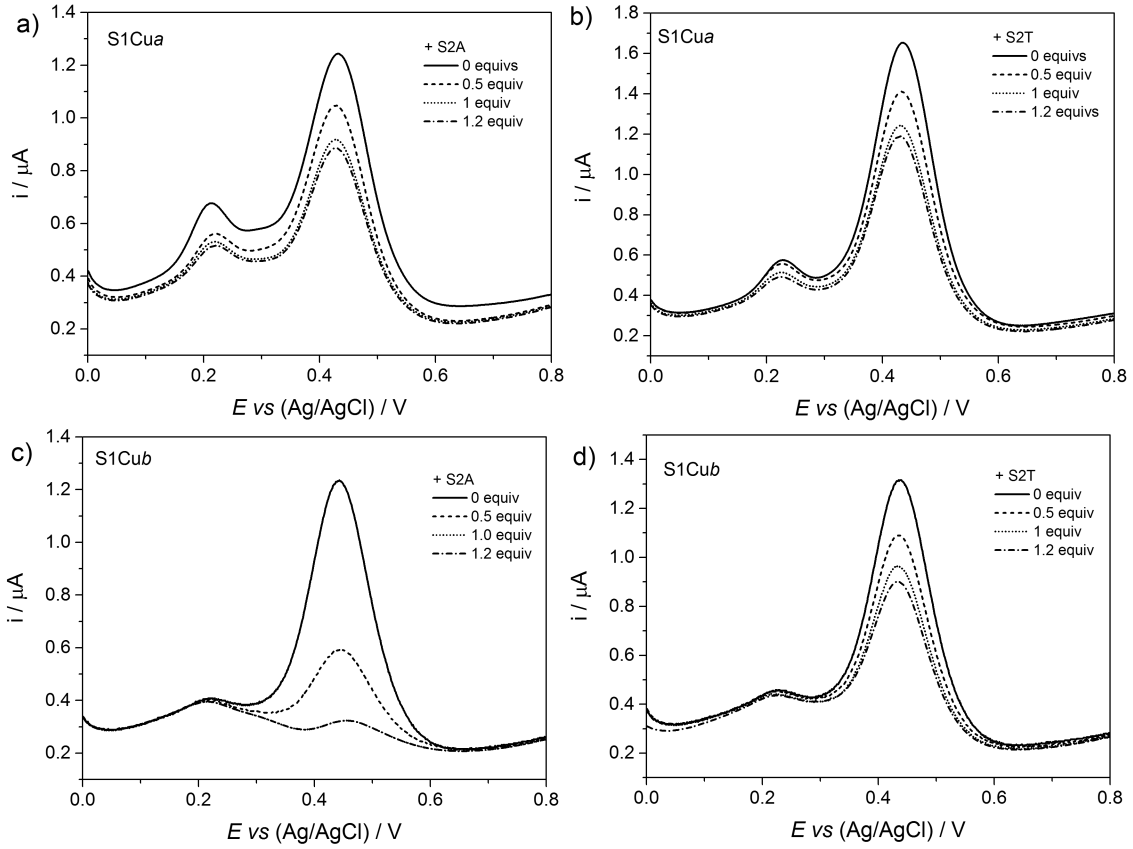
Appendix 8-15 – Examples of analytical HPLC traces for macrocycle modified oligonucleotides. a) ***S1Cua*** *b)* ***S1Cub.***



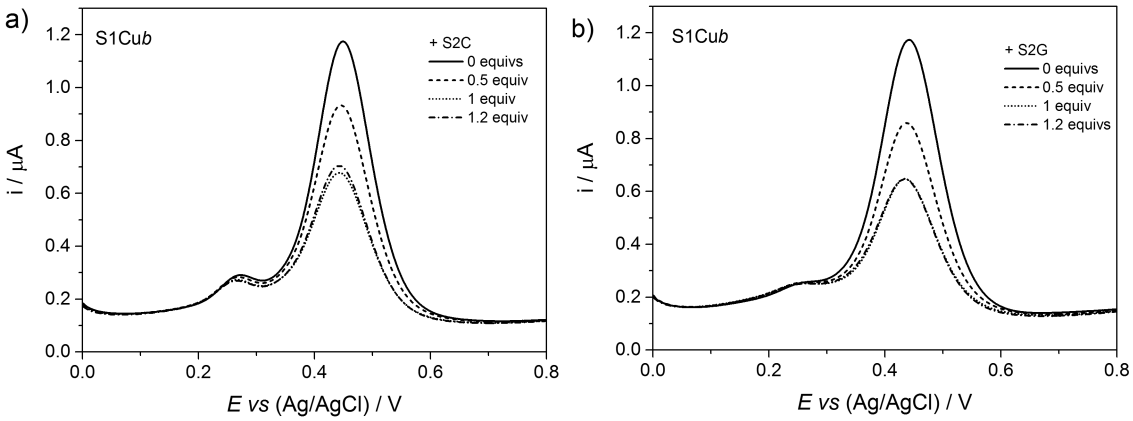
Appendix 8-16 Semi-preparative HPLC trace showing the single peak of interest. Intercalating motifs in comparison (Ni/Cua/b) produce two peaks for each diastereomer.



*Appendix 8-17 - UV-vis titrations of **S1Nia** with a) **S2T** and **S1Nib** with a) **S2A** and b) **S2T** indicating the red shift and decrease in Soret band. Spectra normalized at 260 nm.*



*Appendix 8-18 - SWV of strands **S1Cua** upon addition of targets a) **S2A** and b) **S2T**, **S1Cub** upon addition of targets c) **S2A** and d) **S2T** prior to background subtraction by Origin as shown in Chapter 5.*



Appendix 8-19 - SWV of S1Cub upon addition of increasing molar equivalents of a) S2C and b) S2G prior to background subtraction using Origin.

8.5 DNA labelled Virus Particles

8.5.1 Oligonucleotide Synthesis

HPLC methods used the following gradients;

- ThiolFerro: A two solvent system using Buffer D and MeCN gradients was used in the following time frame;
    - 5-20% MeCN: 35 minutes, 100% MeCN: 7 minutes, 5% MeCN: 8 minutes. Solvent rate: 3 ml per minute.
- Oligo60: A three solvent system using Buffer A, Buffer D and MeCN gradients was used in the following time frames;
    - 30-50% Buffer B: 25 minutes, 50-100% Buffer B: 7 minutes, 100-0% Buffer B: 13 minutes, 0% Buffer B: 10 minutes, 30% Buffer B, 5 minutes
    - 0% MeCN: 35 minutes, 0-100% MeCN: 10 minutes, 100% MeCN: 10 minutes, 0% MeCN: 5 minutes.

Appendix 8-20 - Synthesised Oligonucleotides for Chapter 6

Code	Sequence	Predicted Mass / Da	Observed Mass / Da	HPLC Method	Retention time / min
R-For	5'-SSC6-(TAMRA-dT)ATGAGTATTCAACATTTC-3'	6781	6782	ThiolFerro	22.08
F-Com	5'-(6-FAM)-GAAATGTTGAATACTCAT-3'	6059	6059	Oligo60	36.33
F-Ran	5'-(6-FAM)-TCATCAGTCAGTCAGTCA-3'	6027	6027	Oligo60	37.05

8.5.2 Conjugation calculation

The overall conjugation yield was calculated using the following calculation:

1. $1.1x10^{-10}$ moles of wtM13 (we assume each particle has 2700 pVIII copies).

   Therefore, $1.1x10^{-10}$ x 2700 = **$2.97x10^{-7}$** moles of pVIII available for conjugation.

2. A 10:1 excess of SMCC is added, it is assumed here that we achieve 100% conjugation of SMCC (this is highly unlikely but facilitates the yield calculations).

   Therefore 2.97 x10-7 pVIII-SMCC sites for conjugation with thiolated DNA.

3. $1.17x10^{-8}$ moles of DNA added.

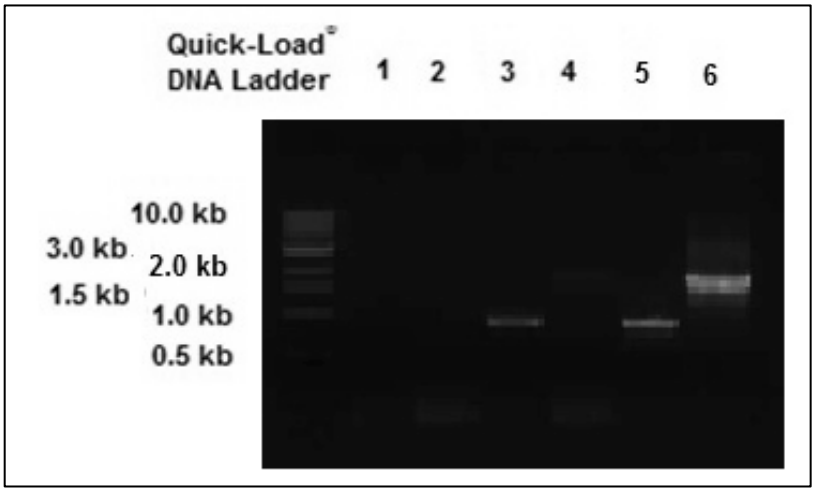
   $((1.17x10^{-8}/2.97x10^{-7})*100)$ = **3.9 %** as the maximum possible coverage i.e. **(106** oligos per particle)

   UV-vis studies (Figure 5) showed the conjugation of approx. 8 oligonucleotides per phage.

   Therefore the overall yield is: (8/106)*100 = **8%**

8.5.3 Use of ethylene glycol

The use of ethylene glycol has been shown to reduce the denaturing temperature of DNA templates for PCR reactions. Appendix 21 below shows how the use of small amounts of ethylene glycol allows the *ampR* gene to be amplified at 86 °C.



Appendix 8-21 -The agarose gel shows the amplicons of the ampR gene following PCR reactions under different conditions. Lane 1) standard reaction at 86 °C (amp.res.), 3) adding ethylene glycol (ampR) 5) standard reaction (positive control for ampR gene).Lanes 2,4 and 5 involve other primer and amplimer products that are not associated with this work.

8.6 List of Publications

The following publications resulted from work carried out during this Ph.D.

Publications:

1. **J. Carr-Smith**, R. Pacheco-Gomez, H. A. Little, M. R. Hicks, S. Sandhu, N. Steinke, D. J. Smith, A. Rodger, S. A. Goodchild, R. A. Lukaszewski, J. H. R. Tucker and T. R. Dafforn, 'Polymerase Chain Reaction on a viral nanoparticle', *ACS Synth. Biol*., 2015, doi: 10.1021/acssynbio.1025b00034.
   **proofs being edited at time of submission.

2. A. Mulas, Y. Willener**, J. Carr-Smith**, K. M. Joly, L. Male, C. J. Moody, S. L. Horswell, H. V. Nguyen and J. H. R. Tucker, 'The effect of central and planar chirality on the electrochemical and chiral sensing properties of ferrocenyl urea H-bonding receptors.' *Dalton Trans.*, 2015, **44**, 7268-7275.
   **this work is not related to the work discussed in this thesis

Oral Presentations:

1. **J. Carr-Smith**, Viral Spaghetti: A Linear Approach to DNA Detection, British Biophysical Society Biennial Meeting 2014, 9th-11th July 2014, Warwick, U.K.

Poster Presentations:

1. **J. Carr-Smith**, R. Pacheco-Gomez, H. A. Little, N. Steinke, J. H. R. Tucker and T. R. Dafforn, 'Linear Dichroism Detection of PCR Amplimer Production Using M13 Bacteriophage', The 11th Nucleic Acid Forum, 4th July 2014, London, U.K.

2. **J. Carr-Smith**, R. Pacheco-Gomez, H. A. Little, N. Steinke, J. H. R. Tucker and T. R. Dafforn, 'Linear Dichroism Detection of PCR Amplimer Production Using M13 Bacteriophage', 3rd Annual Directed Assembly Network Bioinspired & Biomimetic Assembly Theme Meeting, 4th & 5th September 2014, Birmingham, U.K.